BI-PARAMETER MAXIMAL MULTILINEAR OPERATORS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Peter M. Luthy

January 2013

BI-PARAMETER MAXIMAL MULTILINEAR OPERATORS

Peter M. Luthy, Ph.D.

Cornell University 2013

It is well-known that estimates for maximal operators and questions of pointwise convergence are strongly connected. In recent years, convergence properties of so-called 'non-conventional ergodic averages' have been studied by a number of authors, including Assani, Austin, Host, Kra, Tao, and so on. In particular, much is known regarding convergence in L^2 of these averages, but little is known about pointwise convergence. In this spirit, we consider the pointwise convergence of a particular ergodic average and study the corresponding maximal trilinear operator (over $\mathbb{R}$, thanks to a transference principle). Lacey in [17] and Demeter, Tao, and Thiele in [8] have studied maximal multilinear operators previously; however, the maximal operator we develop has a novel bi-parameter structure which has not been previously encountered and cannot be estimated using their techniques. We will carve this bi-parameter maximal multilinear operator using a certain Taylor series and produce non-trivial Hölder-type estimates for the "main" terms by treating them as singular integrals whose symbol's singular set is similar to that of the Biest operator, studied by Muscalu, Tao, and Thiele in [26] and [27]. Modulo further work to estimate certain error terms coming from the Taylor series which a priori seem to be well-behaved, this will allow us to estimate the full bi-parameter maximal multilinear operator.

BIOGRAPHICAL SKETCH

Peter was born in Portland, Maine in August of 1983 to parents Jim Luthy and Nancy Merrow and sisters Alison Fiser (now Ouelett) and Vivian Page; a younger brother, Kurt Luthy, was added just a few years later. Peter spent most of his school years in Portland, though he moved several times during elementary school. Apparently his ability in mathematics developed early, but this interest was not particularly encouraged during school: while being readily able to understand *how* to solve a problem, he was never very good at getting the right answer or at performing computations quickly — the two staples of mainstream mathematical achievement. Peter graduated from Portland High School in 2001, going on to attend Connecticut College in New London, Connecticut, where he studied mathematics and physics. During his third semester, Peter took a complex analysis course with Chikako Mese, now at Johns Hopkins University; near the end of the course, Chika told Peter she thought his proofs were elegant and clever and that he should probably go to graduate school in mathematics. These first true words of encouragement were taken to heart. Peter took many more mathematics courses and attended an REU at Cornell University with Alex Meadows in the Summer of 2004. Peter graduated Summa Cum Laude with distinction in mathematics and physics and won the Oakes and Louise Ames Prize for his honors thesis in mathematics — the first time in the College's history that the award had been offered to mathematics student. In 2005 he began graduate school at Cornell University. In 2007, Peter began working with Camil Muscalu. Working through personal tragedy with Camil's encouragement, Peter was able to eventually finish his degree, albeit a bit slower than he would have liked. After completing his degree in August, Peter will begin his professional mathematical career at Washington University in St. Louis as a William Chauvenet Postdoctoral Lecturer.

I dedicate this work to my father, Jim Luthy, who passed away in June 2009. Though he left sooner than I would have liked, he stayed long enough to leave me in profound gratitude for encouraging and supporting me throughout his life.

ACKNOWLEDGEMENTS

One of my professors here at Cornell, Len Gross, has referred to the process of earning a PhD as something like the process of being born. In the spirit of such a remark, I am reminded of the adage, “It takes a village to raise a child.” As such, I am indebted to a great number of individuals and will take the present opportunity to thank many of them (chronologically, appropriately enough).

First and foremost, I must thank my wonderful family. My mother, Nancy Merrow, and late father, Jim Luthy, did quite a bit more than lend me some excellent genetic material; they also encouraged my sense of curiosity and were committed to developing my education — kicking and screaming, as the case may often have been. My brother, Kurt Luthy, and my sisters, Alison Ouellet and Vivian Page, have been an integral part of my life, and, for some reason, continue to be supportive, despite possibly having been the target of a decent proportion of the aforementioned *alleged* kicking and screaming. I am deeply thankful for all the Vientiane visits, fireman’s carries, and messages from the boss. In addition to my immediate family, there are my sisters’ families: my nieces Iris and Anya and my nephew Julian — all of whom are spectacular tiny humans — and their respective fathers, Justin Deri and Mike Ouellet — who are also quite spectacular, if a bit larger. My cousins Mitchell and Nolan, my Aunt Gerri, and myself all share a certain family resemblance; I learned mine by following their example and also by wearing their hand-me-downs for most of my childhood.

I am profoundly grateful for Jacob Johnson, Chuck Galli, Shoa Clarke, and Brandon Treadwell, all of whom have been wonderful friends for many years and helped me through awkward ages, bad haircuts, and plenty of trips to 7-Eleven. I should thank several other Maine friends as well: Matt Stevenson, Renee Thibodeau, Seth Jewett, Matt “Goody” O’Connor, Kristin Caler Hernandez, Dave

Wagabaza, and Gerald Walker.

I'm indebted to my friends, classmates, and colleagues from Connecticut College and Cornell University who have been there for me in school and beyond. There are far too many for me to thank all of them here, but I'd like to especially mention (in semi-chronological order): Doug Jackson, Tim Gray, Zach West, Ben Giffen, Eli Trowbridge, Amanda Cook, Alex Lanstein, Jeff Langevin, Lauren Eichler, John Mancarella, Chris Mastellone, "Tall" Peter Samuelson, Jason Anema, Sergio Pulido Niño, Chris Scheper, Kristin Oberheide, and Nikolas Rathert. I would also like to single out my adviser's other students from my time at Cornell, Eyvindur Palsson, Joeun Jung, and Cristina Benea, all of whom have been great friends and colleagues.

I would like to take the opportunity to express my gratitude for the many wonderful teachers I have had throughout my life outside of Cornell, especially Portland High School teachers Barbara Solomon, Garry Fox, Stephanie Trainor, Ericka Lee, Toni Skillings, and Scott Wibby; mentors Louis Neveux and Glenn Palomaki from the Foundation for Blood Research; Connecticut College mathematics professors Perry Susskind, Chika Mese, and Christopher Hammond; Connecticut College physics professors Mike Monce, Leslie Brown, and Arlan Mantz; 2004 Cornell REU mentor Alex Meadows; and director of the UCLA Analysis Summer School, Christoph Thiele.

Cornell has been a wonderful place to study mathematics, and I am grateful for all the support of the mathematics department's numerous staff members, past and present, as well as the mathematics faculty. In particular, I'd like to thank Len Gross as well as my committee members Bob Strichartz and Laurent Saloff-Coste for so many educational classes, lectures, and conversations (mathematical and otherwise) — and for inviting a variety of interesting speakers to the analysis seminar. I feel this is an appropriate location to praise Bob Strichartz for establishing and fostering such a wonderful REU program here at Cornell (which I attended in 2004).

Finally, Camil Muscalu has been a truly wonderful adviser. I always leave his office tremendously intellectually inspired. Camil supported and motivated me during the dark days of my father's illness and passing when I found myself unable to concentrate on mathematics. He was open about the loss of his own father and shared some wisdom and advice. Without his encouragement, patience, and understanding, these bleak times would likely have overcome me, and I would probably never have written the document before you. We also had many refreshing conversations outside of mathematics — they truly ran the gamut, from Dennis Rodman's Hall of Fame acceptance speech to living in Boston to Romanian language. Thank you so much, Camil.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 0

INTRODUCTION

0.1 Pointwise Convergence and Maximal Operators

Given a sequence of functions, there is a variety of ways the sequence might converge: pointwise, in norm, weakly, and so on. Pointwise convergence is naïvely the most "natural" but is difficult to work with in the framework of modern analysis. With this in mind, we recall two classical theorems.

Theorem 0.1.1 (Lebesgue Differentiation Theorem)**.** *If $B_r(x)$ denotes the ball of radius r around x in $\mathbb{R}^d$, then given $f \in L^p(\mathbb{R}^d)$ for $p \geq 1$, the average value of f on $B_r(x)$ converges for a.e. x to $f(x)$ as $r \to 0$.*

Theorem 0.1.2 (Carleson–Hunt)**.** *If $\mathbb{T}$ denotes the unit circle and $f \in L^p(\mathbb{T})$ for $p > 1$, then the (symmetric) partial sums of the Fourier series for f converge pointwise to f almost everywhere.*

The proofs of the two theorems are very different, but both come down to proving the theorem for a dense class of functions and that a maximal operator is bounded. Smooth functions serve as suitable dense function classes for both theorems. For Theorem 0.1.1, relevant maximal operator is the well-known Hardy–Littlewood maximal operator M,

$$M(f)(x) = \sup_{r>0} \frac{1}{m(B_r)} \int_{B_r} |f(x+t)| dt,$$

which is bounded from $L^p \to L^p$ for $p \in (1, \infty]$ and bounded from L^1 to weak-L^1; for Theorem 0.1.2, this is the Carleson operator C,

$$C(f)(x) = \sup_{N \in \mathbb{R}} \left| \int_{-\infty}^{N} \hat{f}(\xi) e^{2\pi i x \xi} d\xi \right|,$$

which is bounded from L^p to L^p for $p \in (1, \infty)$. That the Carleson-Hunt theorem is false for $p = 1$ is a result of Kolmogorov from the 1920s and is reflected in the fact that C does not satisfy a suitable L^1 estimate. Hence proving estimates for maximal operators seems to be a main ingredient in proving pointwise convergence theorems. A theorem of Stein says that this is fundamentally true:

Theorem 0.1.3 (Stein, 1961[1]). *Suppose that T_n is a family of bounded linear operators on $L^p(\mathbb{T})$ for $p \in [1, 2]$ which commute with translations (i.e. rotations of the circle). Further, suppose that for each $f \in L^p$ and almost every x, $T_n(f)(x)$ converges pointwise. Then the operator $f \mapsto \sup_n |T_n f|$ is bounded from L^p to weak-L^p.*

Thus there is, to a certain extent, an equivalence of pointwise convergence and boundedness of certain operators, at least in the linear setting.

0.2 Pointwise Convergence in Ergodic Theory

We begin with the following standard definition.

Definition 0.2.1 (Ergodic Transformation). *Let (X, σ, p) be a complete probability space and $T : X \to X$ be an invertible, bimeasurable map which preserves measure, i.e. $pT^{-1}(E) = p(E)$. Let $\mathcal{I}$ denote the collection of sets E with $T^{-1}(E) = E$. $\mathcal{I}$ is called the invariant sigma algebra of T. If $\mathcal{I}$ is the trivial sigma algebra (i.e. every element of $\mathcal{I}$ has probability 1 or 0) we say that T is ergodic.*

Let (X, σ, p) be a complete probability space and suppose that $T : X \to X$ is an invertible, bimeasurable map which preserves measure. If $f \in L^p(X)$, the

[1] This theorem is true in much greater generality, but the requirement that $p \leq 2$ cannot be dropped, in general. See [31, Theorem 1] for the exact statement.

following equality holds almost everywhere:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} f(T^k x) = E(f|\mathcal{I}),$$

where $E(f|\mathcal{I})$ is the conditional expectation of f with respect to the invariant σ-algebra of T. If T is an ergodic transformation, $\mathcal{I}$ is trivial, and so the right side is actually $\int_X f$. In this case, the above equality is the celebrated Birkhoff Ergodic Theorem.

Limits of ergodic averages in the spirit of Birkhoff's theorem have been studied by many authors with a host of applications in mathematics as well as the natural sciences. One of the heralded applications of ergodic theory is Furstenberg's proof of Szemerédi's theorem:

Theorem 0.2.2 (Szemerédi's Theorem)**.** *Any subset of the natural numbers having positive upper density*[2] *contains arithmetic sequences of arbitrary length.*[3]

The main ingredient in Furstenberg's proof is:

Theorem 0.2.3 (Furstenberg's Multiple Recurrence Theorem)**.** *Let* (X, σ, p) *be a probability space and* T *as in Birkhoff's theorem. If* E *has positive measure, then for any* $k > 0$ *there exists an* n *such that*

$$p(E \cap T^{-n}E \cap T^{-2n}E \cap ... \cap T^{-kn}E) > 0.$$

Insofar as Szemerédi's theorem is concerned one should think of T as $T(x) = x + 1$, so that the positivity of the above probability guarantees that E contains some arithmetic sequence of length k. This is not exactly correct — the upper

[2]Here, the upper density of a subset E of the integers is $\lim_{n\to\infty} |[1, n] \cap E| / n$.

[3]Of course this theorem was recently extended to the set of primes by Green and Tao in [14]. This required different methods since the primes do not have positive upper density — by the Prime Number Theorem, the relevant quantity for upper density decays like $1/\log n$.

density is not a probability on $\mathbb{Z}$, for instance — but Furstenberg was able to avoid this technical difficulty. Although Furstenberg's proof avoids the issue, it would be nice if Birkhoff's theorem extended to sequences such as

$$\frac{1}{n}\sum_{k=1}^{n} f_1(T^k x) f_2(T^{2k}x)...f_m(T^{mk}x)$$

converging pointwise to something positive for any m (here one should think that, for all i, $f_i = 1_E$ for some fixed set of positive upper density). For $m = 1$, this is Birkhoff's theorem. The case $m = 2$ was established for $f_1, f_2 \in L^\infty$ by Bourgain, [5], more than twenty years ago. However even for $m = 3$, the question of pointwise convergence of such averages remains open.[4] Recent work of Austin, [2], establishes, among much more general types of averages, that

$$\frac{1}{|I_N|}\sum_{n\in I_N+a_N}\prod_{i=1}^{d} f_i(T^{in}x)$$

converges in L^2-norm to some function whenever I_N is some Følner sequence of subsets of integers — this work generalizes a variety of papers by other authors, e.g. Tao, [33], Host and Kra, [16], and Ziegler, [35]. In the work of Furstenberg and Weiss, [12], expressions like

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^n x) f_2(T^{n^2}x)$$

are also shown to converge in L^2. More complicated averages involving k independent parameters in the sum and $2^k - 1$ functions, such as

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N} f_1(T^n x) f_2(T^m x) f_3(T^p x) f_4(T^{n+m}x) f_5(T^{n+p}x) f_6(T^{m+p}x) f_7(T^{n+m+p}x),$$

are shown to converge almost everywhere by Assani [1].

[4] If one treats the related maximal trilinear operator as a singular integral operator using the methods we will discuss later on, then the related singular integral operator is "morally" the trilinear Hilbert transform, for which no estimates are known.

This large body of work suggested a natural extension, namely whether the bi-parameter average

$$\frac{1}{2M+1}\frac{1}{2N+1}\sum_{m=-M}^{M}\sum_{n=-N}^{N} f_1(T^m x) f_2(T^{-m-n}x) f_3(T^n x) \tag{0.2.1}$$

converges pointwise almost everywhere, where M and N go to infinity at different rates. As discussed in the previous section, questions of pointwise convergence are deeply related to boundedness of maximal operators. Rather than work in the generality of a dynamical system, one can use a correspondence principle to translate the problem to $\mathbb{R}$. For example, see Section 14 of [8]. The maximal operator one produces via such a correspondence principle is precisely

$$(f_1, f_2, f_3) \mapsto \sup_{h_1,h_2} \frac{1}{h_1 h_2}\int_{-h_1}^{h_1}\int_{-h_2}^{h_2} |f_1(x-t) f_2(x+s+t) f_3(x-s)| ds dt. \tag{0.2.2}$$

Forcing $h_1 := h_2$, one obtains, essentially, the object of the main result in [8] by Demeter, Tao, and Thiele. However, the above maximal operator depends on two independent parameters, h_1 and h_2, and so we call it a *bi-parameter maximal operator*. Producing boundedness estimates for this operator will occupy the bulk of this thesis. This is, as far as the author knows, the first time such an operator has been studied.

In particular, we will show that the operator in (0.2.2) is bounded from $L^{p_1} \times L^{p_2} \times L^{p_3} \to L^{p_4}$, for $p_1, p_2, p_3 > 1$ with $1/p_1 + 1/p_2 + 1/p_3 = 1/p_4$, for a "non-trivial" range of exponents p_i.[5] The term "non-trivial" here requires some explanation. One could, for example, assume that $f_2 \in L^\infty$ in which case (0.2.2) splits into a tensor product of Hardy–Littlewood operators and thus Hölder's inequality and well-known results produce "trivial" estimates. However, one would ideally like all the f_i to be as close to L^1 as possible, in which case a number

[5] Clearly, one expects a Hölder-type condition on the exponents since this operator behaves like a pointwise product for a fixed pair h_1, h_2.

of things go awry. Indeed, in such a case, the target space $L^{p_4'}$ has $p_4' < 1$ and $p_4 < 0$, in which case the triangle inequality no longer holds, the relationship between an operator and its adjoint is more complicated, and the 4-linear form one produces by dualizing cannot support Hölder's inequality. Alternatively, one could put $f_3 \in L^\infty$ and invoke other known results — this produces, essentially, a maximal variant of $B(f_1, M(f_2))$, where M is the Hardy–Littlewood maximal operator and B is the bilinear Hilbert transform, an operator which can be handled by the techniques of [17] and [8]. There are a number of such possible trivial estimates which are available. One can then invoke multilinear interpolation results to produce a large family of estimates which require only known results. In this thesis, we produce results outside these easily available estimates to push the range of allowable exponents even further.

0.3 Connection to Singular Integral Operators

Returning to the boundedness of the Hardy-Littlewood maximal operator, we recall that the proof depends on a classical Vitali covering argument; in particular it does not require any Fourier analysis. However, the proof does not extend to the bilinear variant,

$$\sup_{r>0} \frac{1}{2r} \int_{-r}^{r} |f(x+t)g(x+2t)|dt. \tag{0.3.1}$$

This maximal operator corresponds to the pointwise convergence problem of Bourgain described above (modulo some details). Of course, one has immediate estimates for the above expression via Hölder's inequality, but one would like, for example, to have both f and g close to L^1, which cannot be handled by Hölder. One can, however, use techniques from singular integrals to get estimates outside the usual Hölder range. For instance, it was known for a long time that

Littlewood-Paley theory could be used to prove the boundedness of the Hardy-Littlewood operator, even though such sophistication was not necessary. More recently, Lacey, [17], estimated the above maximal operator using methods related to estimating a maximal variant of the bilinear Hilbert transform,

$$\sup_{h>0} \int_{h<|t|<1/h} f(x+t)g(x+2t)\frac{dt}{t}.$$

The results of work by Demeter, Tao, and Thiele, [8], extended this idea to one-parameter maximal n-linear operators by realizing that the n-linear problem is treatable using the techniques from the maximal *bilinear* Hilbert transform. The $m = 3$ generalization of Bourgain's theorem relates to the trilinear Hilbert transform, an operator whose estimation remains an open problem.

The main results of this thesis center on extending these results to bi-parameter maximal operators using the techniques of bi-parameter singular integral operators. In particular, if the Demeter-Tao-Thiele theorem, [8], shows a connection between maximal one-parameter multilinear operators and the maximal bilinear Hilbert transform, the theorem contained within this thesis establishes a connection between bi-parameter maximal multilinear operators and a (doubly) maximal variant of the so-called Biest operator (see [26],[27]) which is connected to AKNS systems — these systems are a way of describing many integrable PDE.

0.4 Biest and AKNS systems

It has been known for some time that there is a strong connection between PDE and time-frequency analysis based on the Heisenberg principle, e.g. as discussed by C. Fefferman in [11]; in this paragraph I will describe a relevant example which inspired the development of the aforementioned Biest operator. In [6] and [7], Christ and Kiselev were interested in proving that eigenfunctions of one-dimensional Schrödinger operators with potential F in L^p are bounded for almost all energies when $p < 2$; in their proof, they produced a collection of multilinear operators T_n and wrote eigenfunctions as a sum of multilinear operators $\sum_n T_n(F, ..., F)$. Their methods broke down when the input functions were all in L^2, although it was conjectured that eigenfunctions would be bounded when $p = 2$. Muscalu, Tao, and Thiele, using time-frequency analysis, showed that some of these multilinear operators were in fact unbounded when the input functions are in L^2 in [25]. This indicates that the multilinear expansion approach is flawed at $p = 2$, though the conjecture may still hold — after all, e^{ix} is a bounded function even though most terms in its power series are not. One can translate the entire discussion to the framework of the aforementioned AKNS systems, to which many integrable PDEs relate. One again produces a family of operators, the simplest of which resemble the Carleson operator and the bilinear Hilbert transform; these are important "protoypical" objects in time-frequency analysis. Indeed, Muscalu, Tao, and Thiele studied a variety of operators arising in this way — the so-called Bi-Carleson, [28], and Biest, [26], [27], operators. Since Muscalu, Tao, and Thiele's approach to the Biest was so fruitful to the results contained in this thesis, we shall present a terse overview of AKNS systems and how they relate to singular integrals.

AKNS[6] systems are systems of ODEs capable of describing a wide variety of PDE, such as the KdV, Nonlinear Schrödinger, and sine-Gordon equations. We proceed with a brief discussion of AKNS systems and the connection to singular integrals. Suppose that $u := (u_1(t), ..., u_n(t))$ is a column vector of complex-valued functions on the line. Let D be a diagonal $n \times n$ matrix with distinct (constant) entries d_i along the diagonal. Suppose that V is a matrix whose entries V_{ij} are

[6]AKNS systems are named after M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur; see, e.g., Chapter 1 in [15]

functions such that diagonal $V_{ii} \equiv 0$. Let λ be a real parameter. One of the defining equations in an AKNS system is

$$\frac{d}{dt}u = i\lambda Du + Vu.$$

The rough (and incorrect) heuristic is that the functions u_i represent the positions in the plane of planets rotating around the origin at rates d_i; the ith planet affects the motion of the jth planet according to the potential V_{ij}.

As a particular example, consider for a fixed function F,

$$\begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = i\lambda \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ F & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

After doing some algebra, one is easily able to produce the time-independent Schrödinger equation,

$$-u_1'' + Fu_1 = \lambda^2 u_1.$$

Going back to the general case: supposing that V is upper-triangular, one has, heuristically, that the mass of each planet is vastly bigger than the next[7]. After a simple substitution, $u_i(t) = w_i(t)e^{id_i t\lambda}$, this equation becomes

$$w' = Vw$$

where $w = (w_1, ..., w_n)$ and $V = (V_{ij}(t)e^{i\lambda(d_i - d_j)t})$. In the simplest case, $n = 2$ and V an upper triangular matrix, one can solve the system exactly to see that $w_2 \equiv C_\lambda$ for some constant C_λ and

$$w_1(t) = C_\lambda \int_{-\infty}^{t} V_{12}(s)e^{i\lambda(d_1 - d_2)s} ds + D_\lambda,$$

[7]For instance with the Sun, Jupiter, and Jupiter's moon Io: each is ≈ 1000 times heavier than the next.

for some constant D_λ. Forgetting the constants and assuming for simplicity $d_1 - d_2 = 1$, we see that bounding $\|w_1\|_\infty$ is equivalent to estimating

$$\sup_{t\in\mathbb{R}} \left| \int_{-\infty}^{t} V_{12}(s) e^{i\lambda s} ds \right|.$$

This is trivially finite if $V_{12} \in L^1$. However, by proving the above expression is p-integrable (with respect to λ) for some p, one immediately gets the expression is finite for almost every λ. By a theorem of Menshov and Zygmund, this is true for $p \in [1, 2)$. Even further, observe that this expression looks very similar to the Carleson operator described at the beginning of this introduction, except that the integrand has V_{12} rather than the Fourier transform thereof. If one presumes that V_{12} is the Fourier transform of some function in L^q for $q \in (1, 2]$, the boundedness of the Carleson operator, along with the Hausdorff-Young inequality, guarantee boundedness of orbits. A similar treatment of the $n = 3$ upper-triangular case produces a maximal bilinear operator, dubbed the Bi-Carleson operator, studied by Muscalu, Tao, and Thiele, [28]. In [26], [27] Muscalu, Tao, and Thiele studied a non-maximal operator, dubbed the Biest, related to the $n = 4$ AKNS. There are certain structural similarities in the form of the Biest operator and the bi-parameter maximal operator studied in this thesis. After transforming the Biest operator into frequency variables, its symbol has discontinuities along two hyperplanes, $\xi_1 = \xi_2$ and $\xi_2 = \xi_3$; one then performs a decomposition with respect to this singular set. It will be convenient to treat our bi-parameter maximal operator in an analogous fashion.

CHAPTER 1

MAIN PROBLEM AND TRANSITION TO DISCRETE MODEL OPERATOR

1.1 Main Problem

For measurable functions f_1, f_2, f_3 with appropriate conditions (to be defined later, but one may assume that these function are smooth, bounded, compactly supported, etc.), our operators, T and T^* are defined by

$$T(f_1, f_2, f_3) = \frac{1}{h_1 h_2} \int_{-h_1}^{h_1} \int_{-h_2}^{h_2} |f_1(x-s)||f_2(x+s+t)||f_3(x-t)| ds dt$$

and

$$T^*(f_1, f_2, f_3) = \sup_{h_1, h_2} \frac{1}{h_1 h_2} \int_{-h_1}^{h_1} \int_{-h_2}^{h_2} |f_1(x-s)||f_2(x+s+t)||f_3(x-t)| ds dt,$$

where the supremum is taken over all real h_1 and h_2. We wish to show that T^* satisfies Hölder-type estimates. Some standard limiting arguments along with restricted weak-type interpolation theorems common in time-frequency analysis will allow us to restrict our attention to smooth functions f_i which are supported on unions of compact intervals such that the f_i have L^∞-norm bounded by 1. We will discuss weak-type interpolation later on. It is often heuristically useful to imagine that the f_i are simply characteristic functions of a union of intervals — the smoothness condition simply makes the Fourier analysis nicer.

A trivial argument shows that it suffices to modify our operator slightly to include only dyadic values of the h_j, i.e. to shift our attention to

$$T^*(f_1, f_2, f_3) = \sup_{k_1, k_2 \in \mathbb{Z}} \frac{1}{2^{k_1} 2^{k_2}} \int_{-2^{k_1}}^{2^{k_1}} \int_{-2^{k_2}}^{2^{k_2}} |f_1(x-s) f_2(x+s+t) f_3(x-t)| ds dt$$

where $k_1, k_2 \in \mathbb{Z}$.

1.1.1 Fourier Representation

In the above, we would like to replace the sharp cutoff functions $\chi_{[-2^{k_1},2^{k_1}]}$ and $\chi_{[-2^{k_2},2^{k_2}]}$ with smooth functions; clearly, it would suffice to replace these sharp cutoffs by Schwartz functions $\theta(2^{-k_i}s)$, say, where θ is non-negative, 1 at 0 and which decays rapidly in units of length 1 away from $[-1,1]$. It may at first glance seem better to pick θ to be compactly supported, but this results in perfect localization in space variables rather than frequency variables. Since we should like to use Fourier analysis, it will be more convenient for the Fourier transforms of the functions be compactly supported. We will define our functions explicitly via the following lemma. First, a definition:

Definition 1.1.1. *For smooth functions η_1, η_2, we define $T^*_{\eta_1,\eta_2}$ as follows:*

$$\sup_{k_1,k_2} \frac{1}{2^{k_1}2^{k_2}} \int_{\mathbb{R}^2} |f_1(x-s)f_2(x+s+t)f_3(x-t)|\check{\eta}_1(s/2^{k_1})\check{\eta}_2(t/2^{k_2})dsdt.$$

Lemma 1.1.2. *There are symmetric, non-negative, real-valued functions α and β which are supported in $[-1,1]$ whose Fourier transforms are non-negative and so that $\check{\alpha}(0) = \check{\beta}(0) = 1$. Moreover,*

$$T^*(f_1,f_2,f_3)(x) \lesssim T^*_{\alpha,\beta}(f_1,f_2,f_3)(x),$$

where the implied constant depends on the choice of α and β.

Proof. Let θ be a nonzero symmetric, real-valued function supported on $[-1/2,1/2]$. Then $\theta * \theta$ is a real-valued symmetric function supported in $[-1,1]$; since θ is symmetric, $\hat{\theta}$ is necessarily real-valued so that $(\hat{\theta})^2 \geq 0$. We may then take α and β to be $(\theta*\theta)^2$, which will again be symmetric, be supported in $[-1,1]$, and have non-negative Fourier transform (being the convolution of non-negative functions); it is also itself non-negative, being the square of a real-valued function.

We also observe that $\widehat{(\theta * \theta)^2}(0) = \int (\theta * \theta)^2(y)dy > 0$, and so we may normalize this function to get $\check{\alpha}(0) = \check{\beta}(0) = 1$.

Since $\alpha, \beta \geq 0$ and $\check{\alpha}(0) = \check{\beta}(0) = 1$, we may choose a constant C, which depends on our choice of θ, so that $\alpha(x/C)\beta(y/C)$ is pointwise greater than $\frac{1}{2}\chi_R$ where R is the rectangle $[-1,1] \times [-1,1]$, which gives the second claim. $\square$

1.1.2 Heuristic: Analogy to Bilinear Hilibert Transform

Ignoring the absolute value signs, we may take the Fourier transform and inverse Fourier transform to produce the following Fourier representation of our operator:

$$\sup_{k_1,k_2} \left| \int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\alpha(2^{k_1}(\xi_1 - \xi_2))\beta(2^{k_2}(\xi_3 - \xi_2))e^{2\pi i x(\xi_1+\xi_2+\xi_3)}d\xi_1 d\xi_2 d\xi_3 \right|,$$

where α and β are of the type given in the previous lemma. It will be more convenient later to reverse the sign of the argument of α, which is harmless, and so we change $\alpha(s)$ to $\alpha(-s)$. Suppose for the moment that α were constant in a small neighborhood of the origin — this is actually impossible since

$$\Delta\alpha(0) = \int \widehat{\Delta\alpha}(\xi)e^{2\pi i \cdot 0 \cdot \xi}d\xi = -(2\pi)^2 \int \xi^2\hat{\alpha}(\xi)d\xi < 0,$$

by the positivity of $\hat{\alpha}$. Ignoring this technical difficulty, we would have that $\alpha(0) - \alpha(s)$ is a function equal to $\alpha(0)$ for $|s| \geq 1$ and 0 in a neighborhood of the origin. The bilinear symbol $\alpha(0) - \alpha(\xi_1 - \xi_2)$ restricted to $\xi_1 < \xi_2$ then looks something like a constant multiple of a scale-truncated Bilinear Hilbert transform — its bilinear symbol is something like $\chi_{\xi_1 < \xi_2}$, and if one broke this function up scale by scale (with respect to the line $\xi_1 = \xi_2$), the α we are now encountering is analogous to a sum of all the scales above 1. Of course we actually have two symbols, $\alpha(\xi_1 - \xi_2)$ and $\beta(\xi_3 - \xi_2)$, which interact with one another. Since the parameters k_1 and

k_2 are independent scale parameters, this gives the impression that our operator corresponds to something like a doubly maximal-variant of two interacting Bilinear Hilbert transforms. Ignoring the maximal nature of such an object, the Biest operator studied by Muscalu, Tao, and Thiele in [26] is of a similar type. Thus there is some hope of borrowing some of their techniques to deal with the present issues.

1.1.3 Making the Analogy Precise

As indicated above, we would prefer if, say, the function α produced in the previous lemma were actually constant in a neighborhood of zero. This is not directly possible. However, we may produce an acceptable substitute via the following technical lemma, which is a slightly modified version of [8, Theorem 3.1]:

Lemma 1.1.3. *Suppose that $\tilde{\alpha}$ and $\tilde{\beta}$ are both constant in $[-1,1]$ and zero outside $[-2,2]$, and*

$$\left|(\tilde{\alpha})^{\vee}(s)\right| \lesssim \frac{1}{(1+|s|)^{M_1}}$$
$$\left|\left(\tilde{\beta}\right)^{\vee}(t)\right| \lesssim \frac{1}{(1+|t|)^{M_2}}.$$

If we can show that $T^*_{\tilde{\alpha},\tilde{\beta}}$ satisfies the desired estimates, depending on M_1, M_2 and on the implied constants in the two inequalities above but not on the particular α, β, then we may pass these estimates to the operators $T^*_{\alpha,\beta}$ above.

Proof. Let $\tilde{\alpha}$ be a smooth, symmetric function which is identically 1 on [-1,1] and supported on $[-2,2]$. Write

$$\alpha(\xi) = \tilde{\alpha}(\xi) + \sum_{u=-\infty}^{0} \phi_u(\xi)$$

where

$$\phi_u(\xi) = (\alpha(\xi) - \tilde{\alpha}(\xi))\left(\tilde{\alpha}(\xi/2^u) - \tilde{\alpha}(\xi/2^{u-1})\right).$$

Perform a similar construction for β using $\tilde{\beta}$ and φ_v. Then by the triangle inequality, we have the following pointwise estimate:

$$T^*_{\alpha,\beta} \lesssim T^*_{\tilde{\alpha},\tilde{\beta}} + \sum_{v=-\infty}^{0} T^*_{\tilde{\alpha},\varphi_v} + \sum_{u=-\infty}^{0} T^*_{\phi_u,\tilde{\beta}} + \sum_{u,v=-\infty}^{0} T^*_{\phi_u,\varphi_v}. \tag{1.1.1}$$

The first term on the right of (1.1.1) obviously satisfies the conditions of the lemma. We now focus on the second term. Observe that φ_v is identically zero on $[-2^v, 2^v]$ and also when $|\xi| \geq 2 \times 2^v$; a similar statement holds for φ_v. Now, since

$$2^v \widehat{\varphi_v(2^v \cdot)}(\xi) = \widehat{\varphi_v}(2^{-v}\xi),$$

it follows that

$$T^*_{\tilde{\alpha},\varphi_v(2^v\cdot)} = T^*_{\tilde{\alpha},\varphi_v},$$

and so

$$T^*_{\tilde{\alpha},\varphi_v} = 2^v T^*_{\tilde{\alpha},\varphi_v(2^v\cdot)/2^v}.$$

Now, we know that $\phi_v(2^v\cdot)/2^v$ is supported inside $[-2,2]$ and is constant on $[-1,1]$. Moreover, we have that

$$\left(\varphi_v(2^v\cdot)\right)^\vee = 2^{-v}\check{\varphi}_v\left(\frac{\xi}{2^v}\right). \tag{1.1.2}$$

By writing $\check{\varphi}_v$ as a convolution and putting the modulus inside the integral from the convolution, it is easy to see that

$$\frac{1}{2^v}\left|\check{\varphi}_v(s)\right| \lesssim 2^v \frac{1}{(1+|s|)^{M_1}}\|\tilde{\alpha}\|_2,$$

where the implied constant depends on α but not v. Plugging this into (1.1.2), we have that

$$\frac{1}{2^v}\left|\left(\varphi_v(2^v\cdot)\right)^\vee(\xi)\right| \lesssim \frac{1}{(1+|s|)^{M_1}}\|\tilde{\alpha}\|_2.$$

This, together with the definition of $\tilde{\alpha}$, guarantees that $T^*_{\tilde{\alpha},\varphi_v(2^v\cdot)/2^v}$ satisfies all the conditions in the statement of the lemma; hence we can translate estimates on $T^*_{\tilde{\alpha},\varphi_v(2^v\cdot)/2^v}$ to

$$\sum_{v=-\infty}^{0} 2^v \left(T^*_{\tilde{\alpha},\varphi_v(2^v\cdot)/2^v}\right) = \sum_{v=-\infty}^{0} T^*_{\tilde{\alpha},\varphi_v},$$

which takes care of the second term on the right of (1.1.1). The last two terms are dealt with in a similar manner. $\square$

The above lemma allows us to assume that the functions α and β appearing in our operator are supported in $[-2,2]$ and constant in $[-1,1]$. In fact, the lemma allows us to assume that they are actually either 0 or 1 in $[-1,1]$. Since we may clearly write such a function which is 0 in $[-1,1]$ as a difference of two functions which are 1 in $[-1,1]$ and 0 outside $[-2,2]$, we make the following assumption:

Assumption 1. *Assume without loss of generality that $\alpha, \beta \equiv 1$ in $[-1,1]$.*

1.2 Discretization.

Because we wish to use Fourier analysis, it is inconvenient (and somewhat pointless) to include the absolute value signs in the above expression. Indeed, by taking the Fourier transform we lose real-ness, let alone positivity. The absolute value of a smooth function need not be smooth but of course can be approximated above by a smooth function. So, we move the absolute value signs outside the integral. After taking the Fourier transform and inverse Fourier transform inside to produce

$$\sup_{k_1,k_2}\left|\int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\alpha(2^{k_1}(\xi_2-\xi_1))\beta(2^{k_2}(\xi_3-\xi_2))e^{2\pi i x(\xi_1+\xi_2+\xi_3)}d\xi_1 d\xi_2 d\xi_3\right|,$$

where α, β satisfy the conditions in Lemma (1.1.3). We take a different approach to that taken in the maximal multilinear paper by Demeter, Tao, and Thiele, [8].

We first use the triangle inequality to consider separately the integrals over each of the four regions of $\mathbb{R}^3$ determined by the two planes $\xi_1 = \xi_2$ and $\xi_2 = \xi_3$ — all four regions are treated identically, so we consider only $\xi_1 < \xi_2 < \xi_3$.

1.2.1 Decomposition Heuristic 1

For a fixed k_1, k_2, our symbol is $\alpha(2^{k_1}(\xi_2-\xi_1))\beta(2^{k_2}(\xi_3-\xi_2))\chi_{\xi_1<\xi_2<\xi_3}$. We will make the following imprecise (and incorrect) observations to get a feeling of what kind of model we should expect. First, in the usual way, one can write $\alpha(2^{k_2}(\xi_2 - \xi_1))$ as a cascading sum of functions $\theta_i(\xi_2 - \xi_1)$ which are supported on bands where $\xi_2 - \xi_1 \approx 2^{-i}$, and likewise for β and functions θ'_j supported on the bands $\xi_3 - \xi_2 \approx 2^{-j}$. We now split the operator into three pieces, namely where $i \gg j$, $i \approx j$, and $j \gg i$, respectively. The piece where $i \approx j$ has only one true scale parameter, and thus the techniques of [8] are, roughly speaking, sufficient. By symmetry, it suffices to consider only $i \gg j$.

For each scale i, one can, heuristically speaking, write $\theta_i(\xi_2 - \xi_1) = \sum_{\ell_1} \phi_i^{\ell_1}(\xi_1)\phi_i^{\ell_1+2}(\xi_2)$, where ϕ_i^s is a function supported in an interval $\omega_{i,\ell_1} := [2^{-i}\ell_1, 2^{-i}(\ell_1 + 1)]$ and are something like the characteristic function of ω_{i,ℓ_1}. This is technically an oversimplification (one truly requires a finite number of expressions involving $\phi_i^{\ell_1}(\xi_1)\phi_i^{\ell_1+n}$, for instance), but we are merely making a heuristic approach anyway, so we ignore these details for the moment. See Figure 1.1 below.

In a similar way, produce functions $\phi_j^{\ell_2}$ for θ'_j. Then one can break our symbol

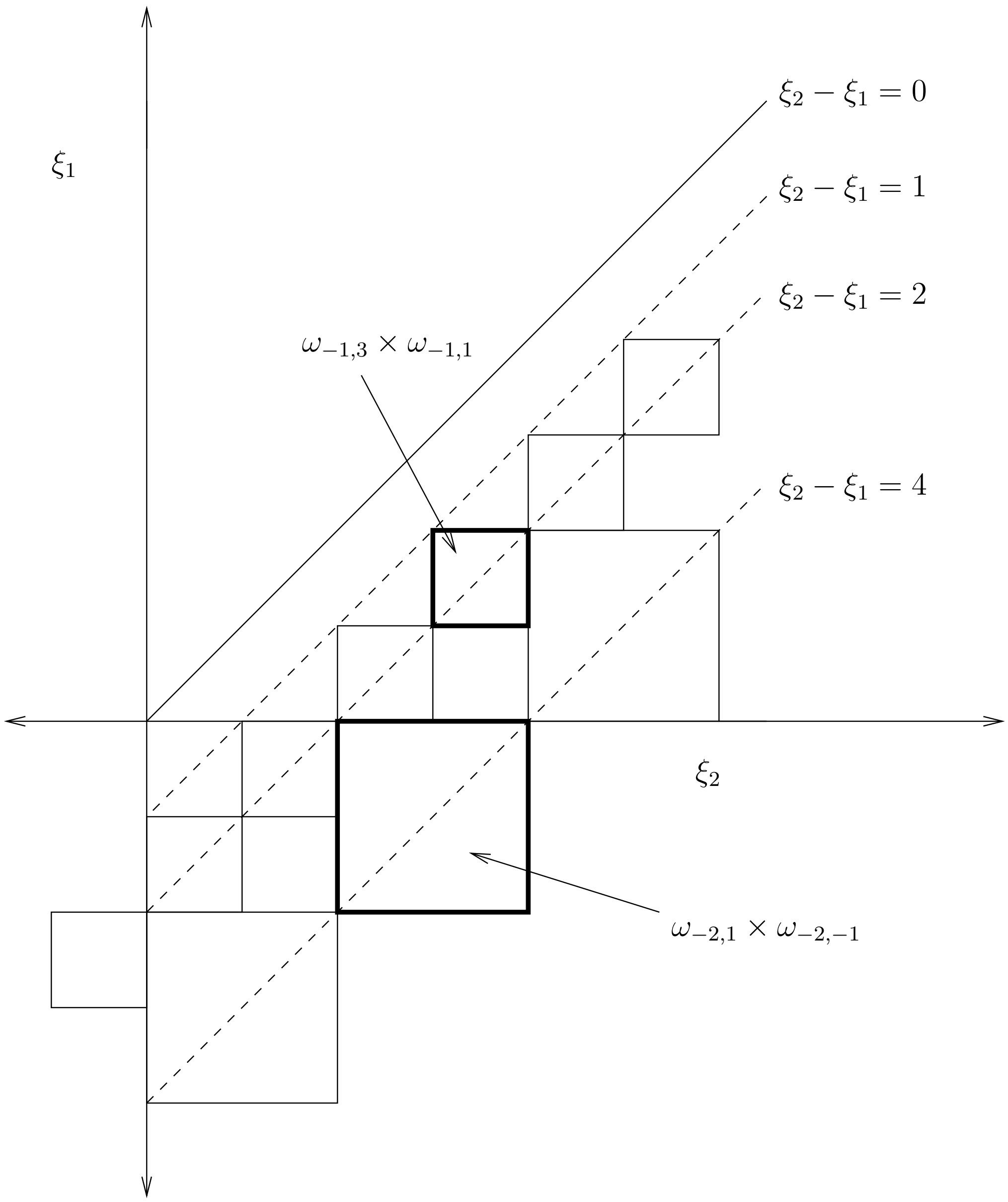

Figure 1.1: A rough visual of how to carve the symbol for $k_2 \gg k_1$.

up as

$$\begin{aligned}
\alpha(2^{k_1}(\xi_1 - \xi_2))\beta(2^{k_2}(\xi_3 - \xi_2)) = &\sum_{i \gg j, i \geq k_1, j \geq k_2} \phi_i^{\ell_1}(\xi_1)\phi_i^{\ell_1+2}(\xi_2)\phi_j^{\ell_2}(\xi_2)\phi_j^{\ell_2+2}(\xi_3) \\
&+ \sum_{i \approx j, i \geq k_1, j \geq k_2} \phi_i^{\ell_1}(\xi_1)\phi_i^{\ell_1+2}(\xi_2)\phi_j^{\ell_2}(\xi_2)\phi_j^{\ell_2+2}(\xi_3) \\
&+ \sum_{i \ll j, i \geq k_1, j \geq k_2} \phi_i^{\ell_1}(\xi_1)\phi_i^{\ell_1+2}(\xi_2)\phi_j^{\ell_2}(\xi_2)\phi_j^{\ell_2+2}(\xi_3).
\end{aligned}$$

Recall that we only consider the $i \gg j$ region. Now, we have that the supports of $\phi_i^{\ell_1+2}(\xi_2)$ and $\phi_j^{\ell_2}(\xi_2)$ must intersect to produce nonzero terms in this sum, and therefore dyadicity of these intervals and the fact that $i \gg j$ guarantees that $\omega_{i,\ell_1+2} \subset \omega_{j,\ell_2}$. We now make another technical oversimplification and presume the following completely false equality: $\phi_i^{\ell_1+2}(\xi_2)\phi_j^{\ell_2}(\xi_2) = \phi_i^{\ell_1+2}(\xi_2)$. This "makes sense" since these functions are to be thought of as characteristic functions and, in any case, the j-function is roughly constant on the interval for the i function by the separation of scales. Then our operator looks like

$$\sum_{i\geq k_1, i\gg j\geq k_2, \omega_{i,\ell_1+2}\subset\omega_{j,\ell_2}} \int_U \widehat{f}_1(\xi_1)\phi_i^{\ell_1}(\xi_1)\widehat{f}_2(\xi_2)\phi_i^{\ell_1+2}(\xi_2)\widehat{f}_3(\xi_3)\phi_j^{\ell_2+2}(\xi_3)e^{2\pi i x(\xi_1+\xi_2+\xi_3)}d\xi,$$

where the integral is over the region

$$U := \{(\xi_1,\xi_2,\xi_3) : \xi_1 < \xi_2 < \xi_3\}.$$

We may now "re-insert" the supremums into our operator and linearize the problem by considering two arbitrary (but fixed) integer-valued functions $N_1(x)$ and $N_2(x)$ to obtain

$$\sum_{i\gg j, \omega_{i,\ell_1+2}\subset\omega_{j,\ell_2}} \int_U \widehat{f}_1(\xi_1)\phi_i^{\ell_1}(\xi_1)\widehat{f}_2(\xi_2)\phi_i^{\ell_1+2}(\xi_2)\widehat{f}_3(\xi_3)\phi_j^{\ell_2+2}(\xi_3)$$
$$\times\, e^{2\pi i x(\xi_1+\xi_2+\xi_3)}d\xi 1_{i\geq N_1(x)}1_{j\geq N_2(x)}.$$

The only caveat is that the estimates must of course be independent of N_1 and N_2. If one now dualizes with a function f_4 and discretizes in the usual way, i.e. as in [22], grouping like scales together, one obtains a model of the form

$$\sum_{i\gg j, m_1\in\mathbb{Z}} |\omega_{i,\ell_1+2}|\langle f_1, \check{\phi}_i^{\ell_1,m_1}\rangle\langle f_2, \check{\phi}_i^{\ell_1+2,m_1}\rangle \left\langle M_{i,\ell_1}(f_3)\check{\phi}_i^{2\ell_1+2,m_1}\chi_{\{|\omega_{i,\ell_1+1}|^{-1}\geq 2^{N_1(x)}\}}, f_4\right\rangle, \tag{1.2.1}$$

where

$$M_{i,\ell_1}(f_3) := \sum_{m_2\in\mathbb{Z}, \omega_{i,\ell_1+2}\subset\omega_{j,\ell_2}} \langle f_3, \check{\phi}_j^{\ell_2,m_2}\rangle\check{\phi}_j^{\ell_2,m_2}\chi_{\{|\omega_{j,\ell_2}|^{-1}\geq 2^{N_2(x)}\}}.$$

Here of course the functions $\check{\phi}_s^{m,n}$ are L^2-normalized functions whose Fourier transforms are supported on intervals of length 2^{-s} translated $\ell_2 \cdot 2^{-s}$ units; moreover, the function itself is "morally" localized to an interval of length 2^s and translated by $m_2 \cdot 2^s$ units.

Ignoring the factor of $M_{i,\ell_1}(f_3)$ — i.e. erasing it completely — one encounters exactly a model of the type found in [8], and their techniques apply directly. The factor M_{i,ℓ_1}, for a fixed m_2, is something like a localized maximal Hilbert transform which depends on the pair i, ℓ_1. One expects, for m_2 very different from the corresponding m_1, quite a bit of decay so that really only the $m_2 \approx m_1$ terms contribute significantly.

The main novelty of the model here is that it has a genuinely bi-parameter structure along with two characteristic functions controlling the scales independently. Thus the techniques of [8] do not apply, and one must obtain new size and energy estimates, which is no small task.

Under the assumption that $i \gg j$ we may invoke the triangle inequality yet again to focus on two separate cases for the supremum: the supremum over k_1, k_2 when $k_1 > k_2$ and when $k_2 \geq k_1$. In the latter case since we have $i \gg j$, we know that $i \gg j > k_2 \geq k_1$, i.e. when $j > k_2$, we automatically have $i > k_1$; thus the supremum can be relaxed to simply a supremum over only k_2 in this case. In the following section, we build the model under the assumption that $k_2 \geq k_1$. The other cases are more delicate and will be written up in later sections.

1.2.2 Taylor series approach for $i \gg j$ and $k_2 \geq k_1$

Since α, β are constant in [-1,1] and supported in [-2,2], we see that

$$\theta(s) = \alpha(s) - \alpha(2s),$$
$$\phi(t) = \beta(t) - \beta(2t),$$

are zero in $[-1, 1]$ and outside of $[-2, 2]$. We now write

$$\theta_i(s) = \theta(2^i s),$$
$$\phi_j(t) = \phi(2^j t).$$

Thus we may write

$$\alpha(2^{k_1} s) = \sum_{i \geq k_1} \theta_i(s),$$
$$\beta(2^{k_2} t) = \sum_{j \geq k_2} \phi_j(t).$$

Hence for any given f_1, f_2, f_3, we may write our maximal operator as

$$\sup_{k_1,k_2} \left| \sum_{i \geq k_1, j \geq k_2} \int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2 - \xi_1)\phi_j(\xi_3 - \xi_2)e^{2\pi i x(\xi_1+\xi_2+\xi_3)} d\vec{\xi} \right|,$$

where $\theta_i(\xi_2 - \xi_1)$ and $\phi_j(\xi_3 - \xi_2)$ are supported in the bands $|\xi_1 - \xi_2| \approx 2^{-i}$ and $|\xi_3 - \xi_2| \approx 2^{-j}$, respectively. As stated previously, we split the interior sum into $i \gg j$, $i \approx j$ and $j \gg i$ and the supremum into the supremum over $k_2 \geq k_1$ and $k_2 < k_1$. More precisely, one may consider the sums where $j > i + 10$, $i > j + 10$ and $|i - j| \leq 10$. Under either assumption that $k_2 \geq k_1$ or $k_2 < k_1$, the restriction to scales where $i \approx j$ is really a finite sum of single-parameter maximal operators nearly identical to those from the work of Demeter, Tao, and Thiele — these operators, after a trivial modification, can all be treated using identical techniques to that of [8]. Thus one only needs to consider the four remaining options, which really consist of two pairs of analogous conditions. Thus it suffices to consider only $i \gg j$ under either the condition $k_2 \geq k_1$ or $k_1 > k_2$.

Assumption 2. *For the remainder of our discussion, we consider only the case $i \gg j$, i.e.*

$$\sup_{k_1,k_2} \left| \sum_{i \gg j, i \geq k_1, j \geq k_2} \int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)\phi_j(\xi_3-\xi_2)e^{2\pi i x(\xi_1+\xi_2+\xi_3)}d\vec{\xi} \right|,$$

where $i \gg j$ means $i > j+10$.

Moreover, as stated in the title of this section, we will focus only on the case when $k_2 \geq k_1$:

Assumption 3. *For the remainder of this section, we discuss only the case $k_2 \geq k_1$ and $i \gg j$, i.e.*

$$\sup_{k_2} \left| \sum_{i \gg j \geq k_2} \int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)\phi_j(\xi_3-\xi_2)e^{2\pi i x(\xi_1+\xi_2+\xi_3)}d\vec{\xi} \right|,$$

where $i \gg j$ means that $i > j+10$.

It will again be convenient to consider the integral only over the set $U \subset \mathbb{R}^3$ where $\xi_1 < \xi_2 < \xi_3$ (the other three analogous regions are treated in the same way, modulo a transposition of indices). In the subset of U where $i \gg j$, we see that any product $\theta_i\phi_j$ is only nonzero in the region $\xi_3 - \xi_2 \gg \xi_2 - \xi_1$ since $\xi_3 - \xi_2 \approx 2^{-j} \gg 2^{-i} \approx \xi_2 - \xi_1$. One of the basic observations from the Biest paper, [27], is that in this region, $\chi_{\xi_1<\xi_2<\xi_3} = \chi_{\xi_1<\xi_2} \cdot \chi_{\xi_1+\xi_2<2\xi_3}$. This latter form is somewhat more convenient: when one discretizes each factor on the right side of this equation, one gets something like $\psi_i^1(\xi_1)\psi_i^2(\xi_2)\psi_j^1(\xi_1+\xi_2)\psi_j^2(\xi_3)$. This is nicer in the sense that the inverse Fourier transform of this is then

$$\left((\check{\psi}_i^1\check{\psi}_i^2) * \check{\psi}_j^1\right) \cdot \check{\psi}_j^2,$$

which is something like a composition of two bilinear Hilbert transforms, where the "inner" BHT is localized to the (larger) frequency interval of the "outer" BHT.

In the Biest paper, [27], Muscalu, Tao, and Thiele are able to subtract from the symbol $\chi_{\xi_1<\xi_2<\xi_3}$ a smooth function which equals $\chi_{\xi_1<\xi_2} \cdot \chi_{\xi_1+\xi_2<2\xi_3}$ in the range $|\xi_3 - \xi_2| \gg |\xi_2 - \xi_1|$ (as well as a second function performing a similar role where $2\xi_1 < \xi_2 + \xi_3$ and $\xi_2 < \xi_3$) to produce something which is a smooth "standard symbol" in that it has only a "nice" singularity along the line $\xi_1 = \xi_2 = \xi_3$ (rather than the two places $\xi_1 = \xi_2$ and $\xi_2 = \xi_3$. We would like to perform a similar move, but this is complicated by the fact that we have something like the symbol for $\chi_{\xi_1<\xi_2<\xi_3}$ which is smoothly *truncated*. When making a similar approach of subtracting "nice" symbols, the fact that this symbol is not identically equal to 1 or 0 has the effect of creating "boundary" terms which are quite complicated, requiring different methods which are apparently as difficult as the ones we presently encounter. We thus abandon the Biest approach in favor of the following methodology. We will still encounter error terms, but they will have a more reasonable shape.

By Taylor's theorem, for a smooth function f,

$$f(x) = f(a) + (x-a)f'(a) + \dots + \frac{(x-a)^n}{n!} f^{(n)}(a) + f_n(x-a),$$

where f_n is the remainder from Taylor's theorem. Thus we may write

$$\phi_j(\xi_3 - \xi_2) = \sum_{m=0}^{n} \frac{(\xi_1 - \xi_2)^m}{2^m m!} \phi_j^{(m)} \left(\xi_3 - \frac{\xi_1 + \xi_2}{2}\right) + \psi_{j,n} \left(\xi_3 - \frac{\xi_1 + \xi_2}{2}\right),$$

where $\psi_{j,n}$ is the remainder term from Taylor's theorem. In particular, by the definition of ϕ_j, it follows that

$$\phi_j^{(m)} \left(\xi_3 - \frac{\xi_1 + \xi_2}{2}\right) = 2^{jm} \phi_{m,j}^* \left(\xi_3 - \frac{\xi_1 + \xi_2}{2}\right),$$

where $\phi_{m,j}^*$ is also a smooth, bounded function supported on the same interval as ϕ_j. Moreover, $(\xi_1 - \xi_2)^m \approx 2^{-im}$ on the support of θ_i, and so $\theta_{i,m}^*(\xi_2 - \xi_1) =$

$2^{im}\theta_i(\xi_2-\xi_1)(\xi_1-\xi_2)^m$ is a bounded, function supported on the same interval as θ_i. Thus for a fixed pair i,j, the m-th order term in the Taylor expansion gains a factor of $2^{-m(i-j)}$, which is small when $i-j$ is big — this holds since we are in the situation that $i \gg j$. We denote by $\tau_{m,k}(\xi_1,\xi_2,\xi_3)$ the symbol which corresponds to the sum of all products $\theta^*_{i,m}\phi^*_{j,m}$ such that $i-j=k\gg 0$ and $j\geq k_2$. Since we are assuming that $i>j+10$, we have that $k>10$. So, the operator whose symbol is the sum of all the m-th order terms is given by $\sum_{k>10}2^{-mk}\tau_{m,k}$. It is not hard to observe that for a finite family of multi-indices α, we may pick m large so that

$$|\partial^\alpha\tau_{m,k}(\xi)|\lesssim 2^{k(m-|\alpha|)}\frac{1}{|\xi|^\alpha},$$

for all α in this family. By doing similar computations for the remainder $\psi_{j,n}$ (and using the remainder theorem for Taylor series), one gets a similar result for the symbol $\tau_{n,k}$ (coming from $\psi_{j,n}$). Thus for sufficiently large n, the $\tau_{n,k}$ satisfy the usual condition for the multilinear Coifman–Meyer multiplier theorem (a recent proof may be found in [21]). We cannot apply the theorem directly, however, since we additionally have a supremum over k_2 still waiting for us. However, this is not a major issue. We will briefly discuss why this is in the following paragraph.

As one can see using the techniques we will use shortly for the $m=0$ term, the discrete model for $\tau_{n,k}$ will be something like

$$\sum_P\langle B_{P,k}(f_1,f_2),\phi^1_P\rangle\langle f_3,\phi^2_P\rangle\langle f_4,\phi^3_P 1_{|I_P|>2^{N_2(x)}}\rangle,$$

where

$$B_{P,k}(f_1,f_2)=\sum_{Q:\omega_{Q_3}\subset\omega_{P_1},\frac{|I_P|}{|I_Q|}=2^k}\langle f_1,\phi^1_Q\rangle\langle f_2,\phi^2_Q\rangle\phi^3_Q.$$

Each interval ω_{P_1} has length 2^{-j} and each ω_{Q_3} has length 2^{-i}. Thus there are precisely 2^k intervals ω_{Q_3} that will contribute to the sum. One can then consider a sum of 2^k models, where the ω_{Q_3} lie in a fixed position within the ω_{P_1} intervals;

if one can estimate each one of these terms separately (in a uniform way), one can estimate the whole model for $\tau_{n,k}$, losing a factor of 2^k in the estimates. As we will discuss, there are sizes and energies available for the $\langle f_3, \phi_P^2\rangle$ term (which is standard) as well as the $\langle f_4, \phi_P^3 1_{|I_P|>2^{N_2(x)}}\rangle$ term (which follows from the methods in [8]). The remaining term, $\langle B_{P,k}(f_1, f_2), \phi_P^1\rangle$ requires a bit more work to estimate fully. However, one can perform some manipulations, provided m is sufficiently large, using some ideas from [20] and [27].

The loss of 2^k is more problematic when $m = 1$ (since we lose a factor of 2^k but only gain a factor of 2^{-k}), but for larger m one will be able to sum over k to get that the full remainder operator, $\sum_{k>10} 2^{-mk}\tau_{n,k}$, is indeed bounded. Thus it suffices to consider the "main term", when $m = 0$, the $m = 1$ term in the Taylor series, and the remainder term from the two-term Taylor expansion.

The Taylor series terms (i.e. when $m = 1$), are, in theory, nicer objects since their symbols have increased in smoothness. Nevertheless, there are some technical issues, and estimating them seems, at present, to require more robust technology than is currently available; thus they will need to be written elsewhere.[1] So, we shall focus only on the $m = 0$ case, as we shall express in the following assumption.

Assumption 4. *For the remainder of this section, we focus on the operator given by the $m = 0$ term in the Taylor expansion described above, i.e. our operator is*

$$\sup_{k_2}\left|\sum_{i\gg j\geq k_2}\int_{\mathbb{R}^3}\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)\phi_j\left(\xi_3-\frac{\xi_1+\xi_2}{2}\right)e^{2\pi ix(\xi_1+\xi_2+\xi_3)}d\vec{\xi}\right|,$$

where $i \gg j$ means $i - j > 10$.

If we dualize with a function f_4, we observe that this last line may be majorized

[1] This issue is precisely the focus of current thesis work by another of Camil Muscalu's students, Joeun Jung.

by

$$\left|\sum_{i\gg j}\int\int_{\mathbb{R}^3}\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)\right.$$
$$\left.\phi_j\left(\xi_3-\frac{\xi_1+\xi_2}{2}\right)e^{2\pi ix(\xi_1+\xi_2+\xi_3)}f(x)1_{j\geq N_2(x)}d\xi dx\right|,$$

for some integer-valued function $N_2(x)$. Thus it suffices to establish estimates for the above which are independent of $N_2(x)$, which we now fix.

Assumption 5. *It suffices to estimate*

$$\left|\sum_{i\gg j}\int\int_{\mathbb{R}^3}\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)\right.$$
$$\left.\phi_j\left(\xi_3-\frac{\xi_1+\xi_2}{2}\right)e^{2\pi ix(\xi_1+\xi_2+\xi_3)}f(x)1_{j\geq N_2(x)}d\xi dx\right|,$$

independent of $N_2(x)$, which is an integer-valued function.

To continue further, we will need to make several standard definitions; we group them together in the following section.

1.2.3 Notation and Definitions

We make the following definitions, which are due to Muscalu, Tao, and Thiele; these statements are copied more or less verbatim from [27, Definitions 4.1–4.6].

Definition 1.2.1. *Let $n \geq 1$ and $\sigma \in \{0, 1/3, 2/3\}^n$. We define the shifted n-dyadic mesh $D = D^n_\sigma$ to be the collection of cubes of the form*

$$D^n_\sigma := \{2^j(k + (0,1)^n + (-1)^j\sigma) : j \in \mathbb{Z} \text{ and } k \in \mathbb{Z}^n\}.$$

We define a shifted dyadic cube to be any member of a shifted n-dyadic mesh.

In our present problem, we will primarily deal with the $n = 3$ case. One can make the standard observation that for any cube Q there exists a shifted dyadic cube Q' such that $Q \subseteq \frac{7}{10}Q'$ and $|Q'| \sim |Q|$.

Definition 1.2.2. *A subset D' of a shifted n-dyadic grid D is called sparse if, for any two cubes Q, Q' in D with $Q \neq Q'$, we have $|Q| < |Q'|$ implies $|10^9 Q| < |Q'|$ and $|Q| = |Q'|$ implies $10^9 Q \cap 10^9 Q' = \emptyset$.*

A standard observation is that any subset of a shifted n-dyadic grid can be split into $O(1)$ sparse subsets.

Definition 1.2.3. *Let $\sigma = (\sigma_1, \sigma_2, \sigma_3) \in \{0, 1/3, 2/3\}^3$, and let $1 \leq i \leq 3$. An i-tile with shift σ_i is a rectangle $I_P \times \omega_P$ with area 1 and with $I_P \in D_0^1$ and $\omega_P \in D_{\sigma_i}^1$. A tri-tile with shift σ is then a 3-tuple $\vec{P} = (P_1, P_2, P_3)$ such that each P_i is an i-tile with shift σ_i and the $I_{P_i} = I_{\vec{P}}$ are independent of i. The frequency cube $Q_{\vec{P}}$ is defined to be $\prod_{i=1}^3 \omega_{P_i}$.*

We shall sometimes abuse notation and refer to i-tiles with shift σ as simply i-tiles or just tiles if it is unimportant or clear from context what the parameters σ and i are.

Definition 1.2.4. *A set $\vec{\mathbf{P}}$ of tri-tiles is called sparse if all tri-tiles in $\vec{P}$ have the same shift and the set $\{Q_{\vec{P}} : \vec{P} \in \vec{\mathbf{P}}\}$ is sparse.*

Clearly by the previous observation, any set of tri-tiles can be split into $O(1)$ sparse subsets.

Definition 1.2.5. *Let P and P' be tiles. We write $P' < P$ if $I_{P'} \subsetneq I_P$ and $3\omega_P \subseteq 3\omega_{P'}$, and $P' \leq P$ if $P' < P$ or $P' = P$. We write $P' \lesssim P$ if $I_{P'} \subseteq I_P$ and $10^7 \omega_P \subseteq 10^7 \omega_{P'}$. We write $P' \lesssim' P$ if $P' \lesssim P$ and $P' \not\leq P$.*

The ordering $<$ is in the spirit of that in Fefferman, [10], or Lacey and Thiele, [18], [19], [34], but slightly different as P' and P do not quite have to intersect. This is more convenient for technical purposes.

Definition 1.2.6. *A collection* $\vec{\mathbf{P}}$ *of tri-tiles is said to have rank 1 if one has the following properties for all* $\vec{P}, \vec{P}' \in \vec{\mathbf{P}}$*:*

1. *If* $\vec{P} \neq \vec{P}'$*, then* $P_j \neq P'_j$ *for all* $j = 1, 2, 3$*.*
2. *If* $P'_j \leq P_j$ *for some* $j = 1, 2, 3$*, then* $P'_i \lesssim P_i$ *for all* $1 \leq i \leq 3$*.*
3. *If in addition to* $P'_j \leq P_j$ *for some* j *we assume that* $|I_{\vec{P}'}| < 10^9 |I_{\vec{P}}|$*, then we have* $P'_i \lesssim' P_i$ *for all* $i \neq j$*.*

Definition 1.2.7. *Let* P *be a tile. A wave packet adapted to* P *is a function* ϕ_P *which has Fourier support in* $\frac{9}{10}\omega_P$ *and obeys the estimates*

$$|\phi_P(x)| \lesssim |I_P|^{-1/2} \tilde{\chi}_I(x)^M$$

for all $M > 0$*, where the implicit constant of course depends on* M *and where*

$$\tilde{\chi}_I(x) := \left(1 + \left(\frac{|x - x_I|}{|I|}\right)^2\right)^{-1/2},$$

where x_I *is the center of the interval* I*.*

1.2.4 Model for $m = 0$ term when $i \gg j$ and $k_2 \geq k_1$, cont.

To reiterate, we are now considering

$$\left| \sum_{i \gg j} \int \int_{\mathbb{R}^3} \hat{f}_1(\xi_1) \hat{f}_2(\xi_2) \hat{f}_3(\xi_3) \theta_i(\xi_2 - \xi_1) \right.$$
$$\left. \phi_j \left(\xi_3 - \frac{\xi_1 + \xi_2}{2}\right) e^{2\pi i x(\xi_1 + \xi_2 + \xi_3)} f(x) 1_{j \geq N_2(x)} d\xi dx \right|.$$

We now proceed through some standard computations. First, we note that $\theta_i(\xi_2 - \xi_1)$ is supported on the set where $\xi_2 - \xi_1 \in [2^{-i}, 2^{-i+1}]$ (recall that we are only considering $\xi_1 < \xi_2 < \xi_3$, and so we ignore the fact that θ_i is actually also nonzero on $[-2^{-i+1}, -2^{-i}]$). We cover this region with a family of shifted dyadic squares, $\mathbf{Q}_\sigma$, where each $Q \in \mathbf{Q}_\sigma$ satisfies $d(Q, \{\xi_1 = \xi_2\}) \approx 2^{-i}$, so that the side length of Q, which we denote $|Q|$, is also approximately 2^{-i-10}. Now produce a family of functions $\psi_{Q,1}(\xi_1), \psi_{Q,2}(\xi_2)$ so that $\psi_{Q,t}$ is supported on $\frac{8}{10}Q_t$ and $\check{\psi}_{Q,t}$ are each adapted to a dyadic interval I_Q (with $|I_Q| = 1/|Q|$) and have $\|\check{\psi}_{Q,t}\|_1 \lesssim 1$. For example, one can construct a function γ which is non-negative and supported on $[0.2, 0.8]$ which decays arbitrarily rapidly away from the origin (since it is necessarily a Schwartz function) and such that

$$\sum_{\ell} \gamma\left(\xi - \frac{\ell}{3}\right) = 1.$$

This is possible because the intervals $[0.2, 0.8]$ translated by multiples of $1/3$ cover the line with enough room for smooth cutoffs. The translation by $\ell/3$ adds a complex exponential to the inverse Fourier transform, which does not affect adaptedness. Thus these functions will suffice. Since we are thinking of these functions as being related to the frequency intervals corresponding to the sides of Q, we will denote these by ω_{Q_1} and ω_{Q_2}, respectively. By these observations, we can choose the $\psi_{Q,t}$ in such a way that

$$a(\xi_1, \xi_2) := \sum_{\sigma \in \{0,1/3,2/3\}^2} \sum_{Q \in \mathbf{Q}_{\sigma,i}} \psi_{Q,1}(\xi_1)\psi_{Q,2}(\xi_2)$$

satisfies

$$a(\xi_1, \xi_2) \equiv 1, \text{ when } \xi_2 - \xi_1 \in [2^{-i}, 2^{-i+1}].$$

Then

$$\theta_i(\xi_2 - \xi_1) = \sum_{\sigma \in \{0,1/3,2/3\}} \sum_{Q \in \mathbf{Q}_{\sigma,i}} \theta_i(\xi_2 - \xi_1)\psi_{Q,1}(\xi_1)\psi_{Q,2}(\xi_2)$$

Let $|Q| := 2^{-i}$. Also, let $\tilde{\psi}_{Q,t}(\xi_t)$ denote a function whose inverse Fourier transform is L^1-normalized and adapted to the same interval I_Q as $\psi_{Q,t}(\xi_t)$ which is 1 on $\frac{8}{10}\omega_{Q_t}$ and 0 outside of $\frac{8.5}{10}\omega_{Q_t}$. Identifying Q with $\mathbb{T}^2$ in the obvious way, we compute a Fourier series to see that

$$\theta_i(\xi_2 - \xi_1)\psi_{Q,1}(\xi_1)\psi_{Q,2}(\xi_2) = \sum_{n_1,n_2} C_1^Q(n_1, n_2)e^{2\pi i \frac{n_1}{|Q|}\xi_1}e^{2\pi i \frac{n_2}{|Q|}\xi_2},$$

on the support of $\tilde{\psi}_{Q,1}(\xi_1)\tilde{\psi}_{Q,2}$. Hence

$$\theta_i(\xi_2 - \xi_1) = \sum_{n_1,n_2} \sum_{\sigma\in\{0,1/3,2/3\}} \sum_{Q\in\mathbf{Q}_{\sigma,i}} C_1^Q(n_1, n_2)\tilde{\psi}_{Q,1}(\xi_1)\tilde{\psi}_{Q,2}(\xi_2).$$

Lemma 1.2.8. *$C_1^Q(n_1, n_2)$ depends only on the σ in the definition of $\mathbf{Q}_{\sigma,i}$ rather than individual i; moreover, it decays arbitrarily rapidly in n_1, n_2. In particular,*

$$|C_1^Q(n_1, n_2)| \lesssim \frac{1}{(1 + |n|)^{M+10}},$$

where M is the decay rate in the definition of a function being adapted to an interval. Lastly, it can be assumed that $C_1^Q(n_1, n_2)$ does not depend on Q, modulo a harmless, finite adjustment of $\mathbf{Q}_{\sigma,i}$ and corresponding finite loss in the estimates. Thus we replace it with $C_1(n_1, n_2)$.

Proof. We see

$$C^Q_{n_1,n_2} = \frac{1}{|Q|^2}\int_{\omega_{Q_1}\times\omega_{Q_2}} Q\theta_i(\xi_2 - \xi_1)\psi_{Q,1}(\xi_1)\psi_{Q,2}(\xi_2)e^{-\frac{2\pi i}{|Q|}(n_1\xi_1+n_2\xi_2)}d\xi_1 d\xi_2.$$

Apply the change of variable $(\xi_1, \xi_2) \mapsto (|Q|\xi_1, |Q|\xi_2)$, one has

$$C^Q_{n_1,n_2} = \int_{I_1\times I_2} \theta(\xi_2 - \xi_1)\psi_{I_1,1}(\xi_1)\psi_{I_2,2}(\xi_2)e^{-2\pi i(n_1\xi_1+n_2\xi_2)}d\xi_1 d\xi_2,$$

where θ lives at scale 1, and the functions $\psi_{I_t,t}(\xi_t)$ live on intervals I_1 and I_2 of scale 1. Moreover, θ is independent of Q. The integral then depends on the difference between the relevant σ_i's involved as well as the distance between centers of the intervals I_1 and I_2 — once one fixes this difference, the integral is always over some rectangle like a fixed $I_1 \times I_2$ except translated parallel to $\xi_1 = \xi_2$, which does not affect the integral. But there are only a finite number of possible distances between the centers (by considering the supports relative to θ, and, modulo a finite loss in the estimates, we may assume the distance is fixed). Repeated applications of integration by parts give the second claim. $\square$

We also write

$$\phi_{Q,t}(\xi_t) := \frac{1}{1+|n_t|^M}\tilde{\psi}_{Q,t}(\xi_t)e^{2\pi i \frac{n_t}{|Q|}\xi_t},$$

and observe the following:

Lemma 1.2.9. *$\phi_{Q,t}(\xi_t)$ is a wave packet adapted to $I_Q \times \omega_{Q_t}$ and has $\|\check{\phi}_{Q,t}\| \lesssim 1$.*

Thus we finally write

$$\theta_i(\xi_2 - \xi_1) = \sum_{n_1,n_2}\sum_{\sigma} C_1(n_1,n_2)(1+|n_1|^M)(1+|n_2|^M)\sum_{Q\in\mathbf{Q}_{\sigma,i}} \phi_{Q,1}(\xi_1)\phi_{Q,2}(\xi_2).$$

It is also clear that for a fixed ξ_1, ξ_2, only finitely many terms in the sum will be nonzero. Performing a similar decomposition to the function $\phi_j(a-b)$ and replacing $a = \xi_3$ and $b = \frac{\xi_1+\xi_2}{2}$, one can write

$$\phi_j\left(\xi_3 - \frac{\xi_1+\xi_2}{2}\right) = \sum_{n_3,n_4}\sum_{\sigma'\in\{0,1/3,2/3\}} \tilde{C}_2(n_3,n_4)\sum_{P\in\mathbf{P}_{\sigma',j}} \phi_{P,1}(\xi_3)\phi_{P,2}\left(\frac{\xi_1+\xi_2}{2}\right),$$

where the $\tilde{C}_2$ has incorporated the polynomial in n_3, n_4 which is present in the previous equation. Hence

Lemma 1.2.10. *Our 4-linear form*

$$\Bigg|\sum_{i\gg j}\int\int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)$$
$$\phi_j\left(\xi_3 - \frac{\xi_1+\xi_2}{2}\right)e^{2\pi i x(\xi_1+\xi_2+\xi_3)}f(x)1_{j\geq N_2(x)}d\xi dx\Bigg|,$$

can be written as

$$\sum_{n\in\mathbb{Z}^4}\sum_{\sigma,\sigma'}C(n)\left|\sum_{i\gg j}\sum_{Q\in\mathbf{Q}_{\sigma,i},P\in\mathbf{P}_{\sigma',j}}\int\int_{\mathbf{R}^3}\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\right.$$
$$\left.\phi_{Q,1}(\xi_1)\phi_{Q,2}(\xi_2)\phi_{P,1}(\xi_3)\phi_{P,2}\left(\xi_1+\xi_2\right)e^{2\pi ix(\xi_1+\xi_2+\xi_3)}f(x)1_{|I_P|\geq 2^{N_2(x)}}d\xi dx\right|,$$

and it suffices to consider fthis or a fixed n *and* σ,σ'*, i.e.*

$$\left|\sum_{i\gg j}\sum_{Q\in\mathbf{Q}_{\sigma,i},P\in\mathbf{P}_{\sigma',j}}\int\int_{\mathbf{R}^3}\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\right.$$
$$\left.\phi_{Q,1}(\xi_1)\phi_{Q,2}(\xi_2)\phi_{P,1}(\xi_3)\phi_{P,2}\left(\frac{\xi_1+\xi_2}{2}\right)e^{2\pi ix(\xi_1+\xi_2+\xi_3)}f(x)1_{|I_P|\geq 2^{N_2(x)}}d\xi dx\right|.$$

Now, since the inverse Fourier transform of

$$\phi_{Q,1}(\xi_1)\phi_{Q,2}(\xi_2)\phi_{P,2}\left(\frac{\xi_1+\xi_2}{2}\right)$$

is

$$\left(\check{\phi}_{Q,1}\check{\phi}_{Q,2}\right)*\check{\phi}_{P,2},$$

it follows that we may insert an L^1-normalized function $\phi_{Q,3}(\xi_1+\xi_2)$ which is 1 on the shifted dyadic interval $\frac{8}{10}\omega_{Q_3}:=\frac{8}{10}(\omega_{Q_1}+\omega_{Q_2})$ and 0 outside $\frac{9}{10}\omega_{Q_3}$. Since $|\omega_{P_2}|\gg|\omega_{Q_3}|$, we must have that $\omega_{Q_3}\subset\omega_{P_2}+\omega_{P_2}:=\omega_{\tilde{P}_2}$ for the product $\phi_{P,2}\phi_{Q,3}$ to be nonzero.

Carrying the inverse Fourier transform through, we produce

$$\left|\sum_{i\gg j}\sum_{P\in\mathbf{P}_{\sigma',j}}\sum_{Q\in\mathbf{Q}_{\sigma,i}:\omega_{Q_3}\subset\omega_{P_2}}\int(f_3*\check{\phi}_{P,1})(x)\right.$$
$$\left.\left((f_1*\check{\phi}_{Q,1})(f_2*\check{\phi}_{Q,2})\right)*\check{\phi}_{Q,3}*\check{\phi}_{P,2}(x)f_4(x)1_{|I_P|\geq 2^{N_2(x)}}dx\right|.$$

One may also insert a factor $\phi_{P,3}$ which is 1 on $\frac{8}{10}\omega_{P_3}:=\frac{8}{10}(\omega_{P_1}+\omega_{P_2})$ and 0 outside

$\frac{9}{10}\omega_{P_3}$, to produce

$$\left|\sum_{i\gg j}\sum_{P\in\mathbf{P}_{\sigma',j}}\sum_{Q\in\mathbf{Q}_{\sigma,i}:\omega_{Q_3}\subset\omega_{P_2}}\int(f_3*\check{\phi}_{P,1})(x)\right.$$
$$\left.\big((f_1*\check{\phi}_{Q,1})(f_2*\check{\phi}_{Q,2})\big)*\check{\phi}_{Q,3}*\check{\phi}_{P,2}(x)\left(f_4 1_{|I_P|\geq 2^{N_2(x)}}\right)*\phi_{P,3}(x)dx\right|.$$

Now, perform a standard discretization procedure with respect to P, as in [30, p. 1654–1656], to produce

$$\left|\int_0^1\sum_{i\gg j}\sum_{P\in\mathbf{P}_{\sigma',j}}\sum_{Q\in\mathbf{Q}_{\sigma,i}:\omega_{Q_3}\subset\omega_{P_2}}\sum_{I_P:|I_P|=|P|^{-1}}\frac{1}{|I_P|^{1/2}}\langle f_3,\phi_{P,1,\alpha}\rangle\right.$$
$$\left.\langle(f_1*\check{\phi}_{Q,1})(f_2*\check{\phi}_{Q,2}))*\check{\phi}_{Q,3},\phi_{P,2,\alpha}\rangle\left\langle f_4 1_{|I_P|\geq 2^{N_2(x)}},\phi_{P,3,\alpha}\right\rangle d\alpha\right|,$$

and perform a second discretization with respect to Q:

$$\left|\int_0^1\int_0^1\sum_{i\gg j}\sum_{P\in\mathbf{P}_{\sigma',j}}\sum_{Q\in\mathbf{Q}_{\sigma,i}:\omega_{Q_3}\subset\omega_{P_2}}\sum_{I_P:|I_P|=|P|^{-1}}\frac{1}{|I_P|^{1/2}}\langle f_3,\phi_{P,1,\alpha}\rangle\right.$$
$$\left\langle\sum_{I_Q:|I_Q|=|Q|^{-1}}\frac{1}{|I_Q|^{1/2}}\langle f_1,\phi_{Q,1,\beta}\rangle\langle f_2,\phi_{Q,2,\beta}\rangle\phi_{Q,3,\beta},\phi_{P,2,\alpha}\right\rangle$$
$$\left.\left\langle f_4 1_{|I_P|\geq 2^{N_2(x)}},\phi_{P,3,\alpha}\right\rangle d\alpha d\beta\right|.$$

Here,

$$\phi_{P,t,\alpha}(x)=|I_P|^{1/2}\overline{\check{\phi}_{P,t}(x-\alpha)}$$

and

$$\phi_{Q,t,\beta}(x)=|I_Q|^{1/2}\overline{\check{\phi}_{Q,t}(x-\beta)}$$

are both L^2-normalized bump functions adapted to the tile $I_P\times P_t$ and $I_Q\times Q_t$, respectively, *uniformly* in α and β. If we let

$$\mathbf{P}:=\{I_P\times P: I_P\text{ dyadic}, |I_P|=2^j, P\in\bigcup_j\mathbf{P}_{\sigma',j}\text{ for some }j\in\mathbb{Z}\}$$

and

$$\mathbf{Q}:=\{I_Q\times Q: I_Q\text{ dyadic}, |I_Q|=2^i, Q\in\bigcup_i\mathbf{Q}_{\sigma,i}\text{ for some }j\in\mathbb{Z}\}$$

then it suffices to study

$$\left|\sum_{P\in\mathbf{P}} \frac{1}{|I_P|^{1/2}} \langle f_3, \phi_{P,1,\alpha}\rangle \left\langle B_P(f_1,f_2), \phi_{P,2,\alpha}\right\rangle \left\langle f_4 1_{|I_P|\geq 2^{N_2(x)}}, \phi_{P,3,\alpha}\right\rangle\right|,$$

where

$$B_P(f_1,f_2) := \sum_{Q\in\mathbf{Q}:\omega_{Q_3}\subset\omega_{P_2}} \frac{1}{|I_Q|^{1/2}} \langle f_1, \phi_{Q,1,\beta}\rangle\langle f_2, \phi_{Q,2,\beta}\rangle \phi_{Q,3,\beta}.$$

Definition 1.2.11. *Let $\vec{\mathbf{P}}$ denote the collection of tri-tiles $\vec{P}$ corresponding to the above construction, and likewise for $\vec{\mathbf{Q}}$.*

Proposition 1.2.12. *Modulo a harmless refinement, the families $\vec{\mathbf{P}}$ and $\vec{\mathbf{Q}}$ are sparse and have rank 1 (see Definition 1.2.6). We may also assume that $\sigma_1 = \sigma_2 = \sigma_1' = \sigma_2' = 0$.*

Proof. With a loss of a factor 3^2, we may assume that $\sigma_1 = \sigma_2$. We also may assume that they are both 0; the other cases are handled precisely the same, modulo some minor changes of notation. Also, by a refinement and loss of $O(1)$ in the estimates, we may freely assume the two families of sparse. We prove the rank 1 condition only for $\vec{\mathbf{P}}$, but the proof works identically for $\vec{\mathbf{Q}}$. We prove each of the three parts of Definition 1.2.6 separately.

1. To establish (1) in the definition, suppose that $P_1 = P_1'$, say. Then clearly the scales of the tiles must be the same; suppose this scale is j. Supposing that the functions $\phi_{P,t}$ live on intervals of slightly smaller scale, say 2^{-j-5}, then by the construction above, if $\xi_1 \in P_1 = [2^{-j-5}\ell_1, 2^{-j-5}(\ell_1+1)]$ and $\xi_2 \in P_2 = [2^{-j-5}\ell_2, 2^{-j-5}(\ell_2+1)]$ then from the fact that $\xi_2 - \xi_1 \in [2^{-j}, 2^{-j+1}]$ (by the factor of $\phi_j(\xi_2-\xi_1)$), it is easy to deduce that $\ell_2 - \ell_1$ can only be selected from a finite family of positive integers (are bounded away from zero as well). Thus we may lose a finite factor in the estimates and assume that

$\ell_2 = \ell_1 + n$ for some fixed positive integer n, which is away from zero. Thus given a P_1, there is exactly one P_2, and hence $P_2 = P_2'$. The definition of P_3 is $P_1 + P_2$, so we know $P_3 = P_3'$ as well. The other two possible cases follow in a similar fashion.

2. Suppose that for some t, $P_t' \leq P_t$. By the previous step, we may assume they are not equal, hence $I_{P'} \subsetneq I_P$ and $3\omega_{P_t} \subset 3\omega_{P_t'}$. Then certainly, $10^7 P_s \lesssim 10^7 P_{s'}$.

3. The P_t intervals are separated by a large number of units of length $|I_P|^{-1}$, and so the third part of the definition holds.

$\square$

By the uniformity of adaptedness in α, β, we may drop the dependence on α, β and will write simply $\phi_P^1 := \phi_{P,1,\alpha}$, since the presence of α does not affect the adaptedness of $\phi_{P,1,\alpha}$ to $I_P \times P_1$. The usual limiting arguments suffice to reduce to finite subsets of $\vec{\mathbf{P}}$ and $\vec{\mathbf{Q}}$.

Assumption 6. *We are now free to study the following for finite families of rank 1 tiles $\vec{\mathbf{P}}$ and $\vec{\mathbf{Q}}$ and functions ϕ_P^t and ϕ_Q^t which are L^2-normalized and adapted in the appropriate way:*

$$\left| \sum_{P \in \vec{\mathbf{P}}} \frac{1}{|I_P|^{1/2}} \langle f_3, \phi_P^1 \rangle \left\langle B_P(f,g), \phi_P^2 \right\rangle \left\langle f_4 1_{|I_P| \geq 2^{N_2(x)}}, \phi_P^3 \right\rangle \right|,$$

where

$$B_P(f,g) := \sum_{Q \in \vec{\mathbf{Q}} : Q_3 \subset \tilde{P}_2} \frac{1}{|I_Q|^{1/2}} \langle f_1, \phi_Q^1 \rangle \langle f_2, \phi_Q^2 \rangle \phi_Q^3,$$

provided the estimates are deduced in a way which does not depend on $\vec{\mathbf{P}}$ and $\vec{\mathbf{Q}}$.

1.3 Taylor series approach when $k_2 \le k_1$.

By way of the same moves as before, we need to consider

$$\left|\sum_{i \gg j} \int \int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2-\xi_1)\right.$$
$$\left.\phi_j\left(\xi_3-\xi_2\right) e^{2\pi i x(\xi_1+\xi_2+\xi_3)} f(x) 1_{i \ge N_1(x)} 1_{j \ge N_2(x)} d\xi dx\right|,$$

for some integer-valued functions $N_1(x)$ and $N_2(x)$, with $N_2(x) \le N_1(x)$, provided our estimates are independent of $N_1(x), N_2(x)$, which we now fix. One cannot proceed precisely as we did before because of the presence of two characteristic functions which associate to the two separate scales. But the general procedure is roughly the same.

As in the previous section, for any pair of scales i, j with $i \gg j$, we may construct a pair of families of cubes $\mathbf{Q}_{i,\sigma}$ and $\mathbf{P}_{j,\sigma'}$ and functions $\psi_{Q,t}, \psi_{P,t}$ and consider

$$\theta_i(\xi_2-\xi_1)\phi_j\left(\xi_3-\xi_2\right)\psi_{Q,1}(\xi_1)\psi_{Q,2}(\xi_2)\psi_{P,2}\left(\xi_2\right)\psi_{P,1}(\xi_3) \tag{1.3.1}$$

Observe that since $i \gg j$, $\omega_{Q_3} := \omega_{Q_1} + \omega_{Q_2}$ is contained in the support of $\psi_{P,2}$ which is ω_{P_2}, so $\omega_{Q_3} \subset \omega_{P_2}$. Now, as before, one inserts a factor

$$\tilde{\psi}_{P,1}(\xi_1)\tilde{\psi}_{P,2}\left(\xi_2\right)\tilde{\psi}_{Q,1}(\xi_1)\tilde{\psi}_{Q,2}(\xi_2),$$

where all these functions are L^1-normalized and $\tilde{\psi}_{Q,t}$ are 1 on $\frac{8}{10}\omega_{Q_t}$ and 0 outside $\frac{8.5}{10}\omega_{Q_t}$, with each adapted, to $I_Q \times \omega_{Q_t}$. The function $\tilde{\psi}_{P,t}$ is 1 on $\frac{8}{10}\omega_{P_t}$ and 0 outside $\frac{8.5}{10}\omega_{P_t}$. One then computes a Fourier series as before.

Thus one has to understand

$$C^{P,Q}(\vec{n})\tilde{\psi}_{P,1}(\xi_3)e^{2\pi i \frac{n_1}{|P|}\xi_3}\tilde{\psi}_{P,2}\left(\xi_2\right)e^{2\pi i \frac{n_2}{2|P|}\xi_2}\tilde{\psi}_{Q,1}(\xi_1)e^{2\pi i \frac{n_3}{|Q|}\xi_1}\tilde{\psi}_{Q,2}(\xi_2)e^{2\pi i \frac{n_4}{|Q|}\xi_2},$$

where, as before, the function $C^{P,Q}(\vec{n})$ decays as rapidly as we'd like in the n_t and, modulo finite refinements, can be made to not depend on P, Q. By the containment of ω_{Q_3} in ω_{P_2}, one has

$$\tilde{\psi}_{P,2}\,(\xi_2)\,\tilde{\psi}_{Q,1}(\xi_1)\tilde{\psi}_{Q,2}(\xi_2) = \tilde{\psi}_{Q,1}(\xi_1)\tilde{\psi}_{Q,2}(\xi_2).$$

By our choice of $\tilde{\psi}_{P,2}$, this changes the above to

$$\begin{aligned} &C^{P,Q}(\vec{n})\tilde{\psi}_{P,1}(\xi_3)e^{2\pi i\frac{n_1}{|P|}\xi_3}e^{2\pi i\frac{n_2}{|P|}\xi_2}\tilde{\psi}_{Q,1}(\xi_1)e^{2\pi i\frac{n_3}{|Q|}\xi_1}\tilde{\psi}_{Q,2}(\xi_2)e^{2\pi i\frac{n_4}{|Q|}\xi_2} \\ &= C^{P,Q}(\vec{n})\tilde{\psi}_{P,1}(\xi_3)e^{2\pi i\frac{n_1}{|P|}\xi_3}\tilde{\psi}_{Q,1}(\xi_1)e^{2\pi i\frac{n_3}{|Q|}\xi_1}\tilde{\psi}_{Q,2}(\xi_2)e^{2\pi i\frac{n_4}{|Q|}\xi_2}e^{\frac{|Q|}{|P|}\pi i\frac{n_2}{|Q|}\xi_2}. \end{aligned}$$

By making use of the decay available, as before we create functions like $\phi_{Q,1} = \tilde{\psi}_{Q,1}(\xi_1)e^{2\pi i\frac{n_3}{|Q|}\xi_1}(1+|n_3|)^{-M}$ and so forth, losing an appropriate factor in the $C^{P,Q}$, which is harmless by the rapid decay (one requires double the decay in the n_2 variable since there are two terms with n_2, but this is not a problem since $C^{P,Q}(n)$ can handle any polynomial loss). These functions will have inverse Fourier transforms which are L^1-normalized and which are adapted to the appropriate tiles. One is still left, of course, with an extra complex exponential.

One finally winds up studying

$$\begin{aligned} &\Bigg|\sum_{\vec{n}\in\mathbb{Z}^4}\sum_{i\gg j}\sum_{Q\in\mathbf{Q}_{\sigma,i}}\sum_{P\in\mathbf{P}_{\sigma',j}:Q_3\subset P_2} C^{P,Q}(\vec{n})\int\int_{\mathbf{R}^3}\hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\phi_{Q,1}(\xi_1) \\ &\phi_{Q,2}(\xi_2)e^{\frac{|Q|}{|P|}\pi i\frac{n_2}{|Q|}\xi_2}\phi_{P,1}(\xi_3)e^{2\pi ix(\xi_1+\xi_2+\xi_3)}f(x)1_{|I_Q|>2^{N_1(x)}}1_{|I_P|\geq 2^{N_2(x)}}d\xi dx\Bigg|. \end{aligned}$$

Now, this is nice but for the factor

$$e^{\frac{|Q|}{|P|}\pi i\frac{n_2}{|Q|}\xi_2}.$$

The function $\psi_{Q,2}$ is supported on a dyadic interval of length $|Q|$ near zero which is then translated by an integer multiple, ℓ_{Q_2}, of $|Q|$. Thus we rewrite the above complex exponential as

$$e^{\frac{|Q|}{|P|}\pi i\frac{n_2}{|Q|}(\xi_2-|Q|\ell_{Q_2})}e^{\frac{|Q|}{|P|}\pi in_2\ell_{Q_2}}.$$

Now, we'd like to write $\exp(\frac{|Q|}{|P|}\pi i \frac{n_2}{|Q|}(\xi_2 - |Q|\ell_{Q_2}))$ as a Taylor series as before, attaching powers of $\xi_2 - \ell_{Q_2}$ to $\phi_{Q,2}(\xi_2)$, and so on. The $m = 0$ case is the one we will focus on, as the higher m will require more robust methods. However, even for $m = 0$, one bumps into an annoying factor $\exp(\frac{|Q|}{|P|}\pi i n_2 \ell_{Q_2})$ which will require additional work since it depends on both Q and P — but when one restricts to a tree, this annoying term is nicely structured. For larger values of m in the Taylor series, one gets essentially the same object as the Taylor series for $k_2 > k_1$, save with a few powers of n_2 (which are harmless thanks to the decay in the function $C^{P,Q}$), and the same number of powers of 2^{-k} where $k = i - j$. We discuss the $m = 0$ case below.

1.3.1 Taylor series approach when $k_2 \leq k_1$ and $m = 0$.

The object we must consider is

$$\Bigg| \sum_{\vec{n}\in\mathbb{Z}^4} \sum_{i \gg j} \sum_{Q\in\mathbf{Q}_{\sigma,i}} \sum_{P\in\mathbf{P}_{\sigma',j}:Q_3\subset P_2} \tilde{C}^{P,Q}(\vec{n}) \int\int_{\mathbf{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\phi_{Q,1}(\xi_1)$$
$$\phi_{Q,2}(\xi_2) e^{\frac{|I_P|}{|I_Q|}\pi i n_2 \ell_{Q_2}} \phi_{P,1}(\xi_3) e^{2\pi i x(\xi_1+\xi_2+\xi_3)} f_4(x) 1_{|I_Q|>2^{N_1(x)}} 1_{|I_P|\geq 2^{N_2(x)}} d\xi dx \Bigg|,$$

where $\tilde{C}^{P,Q}$ is slightly modified from the previous version, but still decays rapidly and can be made to not depend on P, Q as before. To simplify the notation, we will write

$$b_{P,Q} := e^{\frac{|I_P|}{|I_Q|}\pi i n_2 \ell_{Q_2}}.$$

We apply the inverse Fourier transform to get

$$\Bigg| \sum_{\vec{n}\in\mathbb{Z}^4} \sum_{i \gg j} \sum_{Q\in\mathbf{Q}_{\sigma,i}} \sum_{P\in\mathbf{P}_{\sigma',j}:Q_3\subset P_2} \tilde{C}^{P,Q}(\vec{n}) \int (f_1 * \phi_{Q,1})(x)$$
$$(f_2 * \phi_{Q,2})(x)(f_3 * \phi_{P,1})(x) f_4(x) b_{P,Q} 1_{|I_Q|>2^{N_1(x)}} 1_{|I_P|\geq 2^{N_2(x)}} dx \Bigg|.$$

As before, we may insert a factor of $\phi_{Q,3}$ which is 1 on $\omega_{Q_3} := \omega_{Q_1} + \omega_{Q_2}$ and 0 outside a slightly larger interval:

$$\Bigg| \sum_{\vec{n} \in \mathbb{Z}^4} \sum_{i \gg j} \sum_{Q \in \mathbf{Q}_{\sigma,i}} \sum_{P \in \mathbf{P}_{\sigma',j} : Q_3 \subset P_2} \tilde{C}^{P,Q}(\vec{n}) \int (f_1 * \phi_{Q,1})(x)$$
$$(f_2 * \phi_{Q,2})(x) \Big((f_3 * \phi_{P,1}) f_4 b_{P,Q} 1_{|I_Q| > 2^{N_1(x)}} 1_{|I_P| \geq 2^{N_2(x)}} \Big) * \phi_{Q,3}(x) dx \Bigg| .$$

Performing a similar iterated discretization procedure as before results in the following model:

$$\left| \sum_{Q \in \vec{\mathbf{Q}}} \frac{1}{|I_Q|^{1/2}} \langle f_1, \phi_Q^1 \rangle \langle f_2, \phi_Q^2 \rangle \left\langle M_Q(f_3) f_4 1_{|I_Q| \geq 2^{N_1(x)}}, \phi_Q^3 \right\rangle \right| ,$$

where

$$M_Q(f_3) := \sum_{P \in \vec{\mathbf{P}} : \omega_{Q_3} \subset \omega_{P_2}} \langle f_3, \phi_P^1 \rangle \phi_P^1 1_{|I_P| \geq 2^{N_2(x)}} b_{P,Q}.$$

The tile sets $\vec{\mathbf{P}}$ and $\vec{\mathbf{Q}}$ can both be taken to be finite, rank 1 sets. The $b_{P,Q}$ here is an added complication, but we shall see that, when restricting to certain collections of tiles called trees, this extra complex exponential will factor into two pieces, one which depends only on the tree and a second which can be expanded in a second Taylor series. These trees will be described in later sections.

1.3.2 Taylor series approach for Taylor remainder (i.e. $m \gg 0$) when $k_2 \leq k_1$.

If $m \gg 0$, then, after performing a few manipulations, one gets, essentially,

$$\Bigg| \sum_{k \geq 10} 2^{-mk} \sum_{\vec{n} \in \mathbb{Z}^4} n_1^m \sum_{i \gg j} \sum_{P \in \mathbf{P}_{\sigma',j}} \sum_{Q \in \mathbf{Q}_{\sigma,i} : Q_3 \subset P_2, |I_Q| = 2^k |I_P|} C^{P,Q}(\vec{n}) \int \int_{\mathbf{R}^3} \hat{f}_1(\xi_1) \hat{f}_2(\xi_2) \hat{f}_3(\xi_3)$$
$$\phi_{Q,1}(\xi_1) \tilde{\phi}_{Q,2}(\xi_2) \phi_{P,1}(\xi_3) e^{2\pi i x(\xi_1 + \xi_2 + \xi_3)} f_4(x) b_{P,Q} 1_{|I_Q| > 2^{N_1(x)}} 1_{|I_P| \geq 2^{N_2(x)}} d\xi dx \Bigg| .$$

It is important to note that since we know precisely the difference $i-j$ for each term, we may change the condition $|I_Q| \geq 2^{N_1(x)}$ with $|I_P| \geq 2^{N_1(x)-k}$. Thus if we abuse notation slightly to produce $N_k(x) := \min\{N_2(x), N_1(x)-k\}$, we may change the above to

$$\Bigg| \sum_{k\geq 10} 2^{-mk} \sum_{\vec{n}\in\mathbb{Z}^4} n_1^m \sum_{i\gg j} \sum_{P\in\mathbf{P}_{\sigma',j}} \sum_{Q\in\mathbf{Q}_{\sigma,i}:Q_3\subset P_2, |I_Q|=2^k|I_P|} C^{P,Q}(\vec{n}) \int\int_{\mathbf{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)$$
$$\phi_{Q,1}(\xi_1)\tilde{\phi}_{Q,2}(\xi_2)\phi_{P,1}(\xi_3)e^{2\pi i x(\xi_1+\xi_2+\xi_3)} f_4(x) b_{P,Q} 1_{|I_P|\geq 2^{N_k(x)}} d\xi dx \Bigg|.$$

Now, one inserts back functions like $\tilde{\psi}_{Q,3}(\xi_1+\xi_2)$ and $\tilde{\psi}_{P_2}(\xi_1+\xi_2)$ and proceeds as we did in the discussion of the Taylor series approach when $k_2 > k_1$. It is more convenient to switch the order of the sum over P and Q; neglecting a few harmless powers of $|\vec{n}|$ (provided one does not go too high in m), one produces, essentially,

$$\left| \sum_{k\geq 10} 2^{-mk} \sum_{P\in\vec{\mathbf{P}}} \frac{1}{|I_P|^{1/2}} \langle f_3, \phi_P^1\rangle \langle \tilde{B}_{P,k}(f_1,f_2), \phi_P^2\rangle \left\langle f_4 1_{|I_P|\geq 2^{N_k(x)}}, \phi_P^3 \right\rangle \right|,$$

where

$$\tilde{B}_{P,k}(f_1,f_2) := \sum_{Q\in\vec{\mathbf{Q}}:\omega_{Q_3}\subset\omega_{P_2}, |I_Q|=2^k|I_P|} b_{P,Q} \langle f_1, \phi_Q^1\rangle \langle f_2, \phi_Q^2\rangle \phi_Q^3,$$

which is nearly the same model as we produced before for the Taylor remainder when $k_2 < k_1$, modulo the extra factor $b_{P,Q}$. Since the size and energy estimates for the third term in the first of the two previous expressions are made independently of the choice of $N_1(x), N_2(x)$, we are able to estimate these terms for every k and then sum them up, provided we accept, as we did before for the Taylor remainder when $k_2 > k_1$, a loss of 2^k. Thus the Taylor remainder term should be bounded, observing that, for a fixed ratio $|I_P|/|I_Q| = 2^{-k}$, we have that $b_{P,Q}$ depends only on Q and still restricts to trees in a nice way (as we shall see for the sizes in the $m=0$ case). However, for the $m=1$ case, for example, more robust methods are required, as before.[2]

[2] In this case, one should again be able to use methods currently being explored by Joeun Jung, another of Camil Muscalu's students, modulo the $b_{P,Q}$ term, which restricts nicely to trees.

CHAPTER 2

SIZES AND ENERGIES FOR $M = 0$, $K_2 > K_1$

Notation 1. *For ease of writing, we will make the following notation:*

$$\tilde{\phi}_P^3 := \phi_P^3 1_{|I_P| \geq 2^{N_2(x)}}$$

We also recall the following:

Notation 2. *Given a rank 1 family of tri-tiles $\vec{\mathbf{Q}}$, suppose that $Q \in \vec{\mathbf{Q}}$. Q is then made up of three tiles, each given by the product of a fixed interval I_Q with a frequency interval, which we will denote ω_{Q_t}, $t = 1, 2, 3$.*

The model in question is given by

$$\sum_{P \in \vec{\mathbf{P}}} \frac{1}{|I_P|^{1/2}} \langle f_3, \phi_P^1 \rangle \left\langle B_P(f_1, f_2), \phi_P^2 \right\rangle \left\langle f_4, \tilde{\phi}_P^3 \right\rangle,$$

where

$$B_P(f_1, f_2) := \sum_{Q \in \vec{\mathbf{Q}} : \omega_{Q_3} \subset \omega_{\tilde{P}_2}} \frac{1}{|I_Q|^{1/2}} \langle f_1, \phi_Q^1 \rangle \langle f_2, \phi_Q^2 \rangle \phi_Q^3,$$

where $\vec{\mathbf{P}}$ and $\vec{\mathbf{Q}}$ are sparse, finite, rank 1 families of tri-tiles.

Following the standard multilinear harmonic analysis approach, as in [8], [27], [22], and many others, we wish to discuss sizes, which will require the notion of a tree.

Definition 2.0.1. *For any $t \in \{1, 2, 3\}$ and a tri-tile $\vec{P}_T \in \vec{\mathbf{P}}$, we define a j-tree with top $\vec{P}_T$ to be a collection of tri-tiles $T \subset \vec{\mathbf{P}}$ such that*

$$P_j \leq P_{T,t} \text{ for all } \vec{P} \in T,$$

where $P_{T,t}$ is the t-component of $\vec{P}_T$. We will write I_T and $\omega_{T,t}$ for $I_{\vec{P}_T}$ and $\omega_{\vec{P}_{T,t}}$, respectively. We say that T is a tree if it is a t-tree for some $1 \leq t \leq 3$.

It is worth remarking that a tree does not necessarily have to contain its top.

Definition 2.0.2. *We will say that a tree T is t-lacunary if it is a t'-tree for some $t \neq t'$.*

Definition 2.0.3. *Let $t \in \{1, 2, 3\}$. Two trees T and T' are said to be strongly i-disjoint if*

1. *$P_i \neq P'_i$ for all $\vec{P} \in T$ and $\vec{P}' \in T'$.*
2. *Whenever $\vec{P} \in T$, $\vec{P}' \in T'$, are such that $2\omega_{P_i} \cap 2\omega_{P'_i} \neq \emptyset$, then one has $I_{\vec{P}'} \cap I_T = \emptyset$, and similarly with T and T' reversed.*

2.1 Sizes

Definition 2.1.1. ***Sizes.*** *Suppose that $\vec{\mathbf{P}}$ is a finite collection of tri-tiles and $t \in \{1, 2, 3\}$. Suppose also that $(a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}$ is a sequence of complex numbers. Here one really should think of a_{P_j} as being a sequence "living" on the tiles P_j rather than the full tri-tile.*

$$\mathrm{size}_j\big((a_{P_j})_{\vec{P} \in \vec{\mathbf{P}}}\big) := \sup_{T \subset \vec{\mathbf{P}}} \left(\frac{1}{|I_T|} \sum_{\vec{P} \in T} |a_{P_j}|^2 \right)^{1/2},$$

where the T in the supremum ranges over all trees in $\vec{\mathbf{P}}$ which are i-trees for some $i \neq j$. In other words, the supremum ranges over all trees which are j-lacunary.

The above definitions work for general sequences, but for our purposes, we should keep in mind that the sequences we are interested in are

1. $a_{P_1} = \langle f_1, \phi_P^1 \rangle$

2. $a_{P_2} = \langle B_P(f,g), \phi_P^2 \rangle$

3. $a_{P_3} = \left\langle f_4, \tilde{\phi}_P^3 \right\rangle$

The heuristic meaning of these sizes is that the size of a sequence is a measure the extent to which it can concentrate on a single tree. It should be thought of as a phase-space variant of the BMO norm. Indeed, one has a relevant variant of the John-Nirenberg inequality:

Proposition 2.1.2. *If $\mathcal{I}$ is a finite family of dyadic intervals, r is any positive real number and $(a_I)_{I \in \mathcal{I}}$, then define $\|(a_I)_I\|_{BMO(r)}$ by*

$$\|(a_I)_I\|_{BMO(r)} := \sup_{I_0 \in \mathcal{I}} \frac{1}{|I_0|^{\frac{1}{r}}} \left\| \left(\sum_{I \subseteq I_0} \frac{|a_I|^2}{|I|} \chi_I(x) \right)^{1/2} \right\|_r .$$

Then if $0 < p < q < \infty$,

$$\|(a_I)_I\|_{BMO(p)} \sim \|(a_I)_I\|_{BMO(q)}.$$

Proof. See the appropriate section of Chapter 2 of [22]. □

The sizes defined above roughly correspond to this $BMO(r)$ norm when $r = 2$. We state several lemmas which will be used to estimate our model. Since we will be using restricted weak-type interpolation (explained later on), we need one more definition:

Definition 2.1.3. *Suppose that E is a set of finite measure. We define the space $X(E)$ to denote the space of all functions f supported on E with $\|f\|_\infty \leq 1$.*

The following three lemmas are the size estimates we require:

Lemma 2.1.4. *Let E_1 be a set of finite measure, let f_3 be in $X(E_3)$, and let $\vec{\mathbf{P}}$ be a finite collection of tri-tiles. Then one has*

$$\mathrm{size}_1((\langle f_3, \phi_P^1\rangle)_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim \sup_{\vec{P}\in\vec{\mathbf{P}}} \frac{\int_{E_3} \tilde{\chi}_{I_{\vec{P}}}^M}{|I_{\vec{P}}|},$$

for all $M > 0$, with the implicit constant depending on M.

Proof. See Lemma 6.8 in [27]. □

Lemma 2.1.5. *Let E_1, E_2 be sets of finite measure, let f_1, f_2 be in $X(E_1)$ and $X(E_2)$, respectively, and let $\vec{\mathbf{P}}$ be a finite collection of tri-tiles. Let*

$$(a_{P_2})_{\vec{P}\in\vec{\mathbf{P}}} := \left(\sum_{Q\in\vec{\mathbf{Q}}:\omega_{Q_3}\subset\omega_{\vec{P}_2}} \frac{1}{|I_Q|^{1/2}}\langle f_1, \phi_Q^1\rangle\langle f_2, \phi_Q^2\rangle\phi_Q^3, \phi_P^2\right)_{\vec{P}\in\vec{\mathbf{P}}}.$$

Then one has

$$\mathrm{size}_2\left((a_{P_2})_{\vec{P}\in\vec{\mathbf{P}}}\right) \lesssim \sup_{\vec{P}\in\vec{\mathbf{P}}} \left(\frac{\int_{E_1} \tilde{\chi}_{I_{\vec{P}}}^M}{|I_{\vec{P}}|}\right)^{\theta} \left(\frac{\int_{E_2} \tilde{\chi}_{I_{\vec{P}}}^M}{|I_{\vec{P}}|}\right)^{1-\theta},$$

for any $0 < \theta < 1$ and all $M > 0$, with the implicit constant depending on θ, M.

Proof. See Lemma 9.1 in [27]. □

Finally, we state Theorem 6.2 from [8], modulo a trivial change of function space from $X_2(E_4)$ to $X(E_4)$. This statement is of the same flavor as the previous two, modulo a technical limitation which seemingly can be done away with, but we leave it in for ease of use:

Lemma 2.1.6. *Let E_4 be a set of finite measure, let f_4 be in $X(E_4)$, and let $\vec{\mathbf{P}}$ be a finite collection of tri-tiles. We define*

$$\mathcal{I}_{\vec{\mathbf{P}}} := \{I \text{ dyadic } : I_{\vec{P}} \subseteq I \subseteq I_{\vec{P}'} \text{ for some } \vec{P}, \vec{P}' \in \vec{\mathbf{P}}\}$$

Then one has

$$\text{size}_3((\langle f_4, \tilde{\phi}_P^3\rangle)_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim \sup_{I\in\mathcal{I}_{\vec{\mathbf{P}}}} \frac{\int_{E_4} \tilde{\chi}_I^M}{|I|},$$

for all $M > 0$, with the implicit constant depending on M.

2.2 Energies

We define the energies in this case as follows. The 1- and 2-energies are modified somewhat from the "standard" energies.

Definition 2.2.1. *If $t = 1$ or $t = 2$*

$$\text{energy}_t((a_{P_t})_{\vec{P}\in\vec{\mathbf{P}}}) := \sup_{n\in\mathbb{Z}} \sup_{\mathcal{F}} 2^n \left(\sum_{T\in\mathcal{F}} |I_T|\right)^{1/2},$$

where the second supremum ranges over all forests $\mathcal{F}$ consisting of strongly t-disjoint t-lacunary trees in $\vec{\mathbf{P}}$ such that

$$\left(\sum_{\vec{P}\in T} |a_{P_t}|^2\right)^{1/2} \geq 2^n |I_T|^{1/2}$$

for all $T \in \mathcal{F}$ and

$$\left(\sum_{\vec{P}\in T'} |a_{P_t}|^2\right)^{1/2} \leq 2^{n+1} |I_{T'}|^{1/2}$$

for all sub-trees $T' \subset T \in \mathcal{F}$.

And here are the relevant estimates for the 1- and 2-energies:

Lemma 2.2.2. *Let f_3 be a function in $X(E_3)$ and $\vec{\mathbf{P}}$ a finite collection of tri-tiles. Then*

$$\text{energy}_1((\langle f_3, \phi_P^1\rangle)_{\vec{P}\in\vec{\mathbf{P}}}) \leq |E_3|^{1/2}.$$

Proof. See Lemma 6.7 from [27]. $\square$

Lemma 2.2.3. *Suppose E_1, E_2 be sets of finite measure and f_1, f_2 functions with $f_1 \in X(E_1)$ and $f_2 \in X(E_2)$. Let*

$$(a_{P_2})_{\vec{P}\in\vec{\mathbf{P}}} := \left(\sum_{Q\in\vec{\mathbf{Q}}:\omega_{Q_3}\subset\omega_{\tilde{P}_2}} \frac{1}{|I_Q|^{1/2}} \langle f_1, \phi_Q^1\rangle\langle f_2, \phi_Q^2\rangle \phi_Q^3, \phi_P^2 \right)_{\vec{P}\in\vec{\mathbf{P}}}.$$

Then

$$\text{energy}_2\left((a_{P_3})_{\vec{P}\in\vec{\mathbf{P}}}\right) \lesssim \left(|E_1|^{1/2} \sup_{\vec{Q}\in\vec{\mathbf{Q}}} \frac{\int_{E_2} \tilde{\chi}_{I_{\vec{Q}}}^M}{|I_{\vec{Q}}|}\right)^{\theta} \left(|E_2|^{1/2} \sup_{\vec{Q}\in\vec{\mathbf{Q}}} \frac{\int_{E_1} \tilde{\chi}_{I_{\vec{Q}}}^M}{|I_{\vec{Q}}|}\right)^{1-\theta}.$$

for any $0 < \theta < 1$

Proof. See Lemma 9.2 in [27], modulo some obvious changes of notation. $\square$

There is not exactly a 3-energy. However, we have the following replacement which is of a similar flavor:

Lemma 2.2.4. *Let $\mu > 0$. Suppose that $\mathcal{F}$ is a forest of strongly 3-disjoint, 3-lacunary trees. Suppose further that $f \in X(E)$ is such that*

$$\left(\sum_{\vec{P}\in T} |\langle f_4, \tilde{\phi}_P^3\rangle|^2\right)^{1/2} \geq 2^n |I_T|^{1/2}$$

for all $T \in \mathcal{F}$ and

$$\left(\sum_{\vec{P}\in T'} |\langle f_4, \tilde{\phi}_P^3\rangle|^2\right)^{1/2} \leq 2^{n+1} |I_{T'}|^{1/2}$$

for all sub-trees $T' \subseteq T \in \mathcal{F}$. Then

$$\left(\sum_{T\in\mathcal{F}} |I_T|\right)^{1/2} \lesssim |E_4|^{1/2} 2^{-n} \left(2^{-n}|E_4|^{-1/2}\right)^{1/\mu},$$

where the implicit constant depends on μ.

Proof. See Lemma 9.2 in [8]. This is the primary lemma of Demeter, Tao, and Thiele's paper and requires roughly 20 pages of computations. The main idea is the following. Let

$$N_{\mathcal{F}} := \sum_{T \in \mathcal{F}} 1_{I_T},$$

and suppose that I_0 is any interval which contains the support of $N_{\mathcal{F}}$. With a lot of hard work and the help of a theorem of Rademacher-Menshov and a lemma of Bourgain, one can establish the following for any $\mu > 0$:

$$\sum_{\vec{P} \in \cup_{T \in \mathcal{F}} T} |\langle f_4, \phi_P^3 1_{|I_P| > 2^{N(x)}} \rangle|^2 \lesssim \|N_{\mathcal{F}}\|_\infty^{1/\mu} \int |f_4|^2 \chi_{I_0}^{10}.$$

More precisely, one shows that one loses at most a small power of the logarithm of $\|N_{\mathcal{F}}\|_\infty$. The two hypotheses guarantee that our estimate is still preserved after restricting to subtrees, which, it turns out, is precisely enough to get the desired conclusion. □

The factor $(2^{-n})^{1/\mu}$ is essentially technical and can basically be ignored; however, its presence bars one from taking the desired supremum over n in the definitions of 1- and 2-energies. That said, we can use this lemma to establish the following:

Lemma 2.2.5. *Let $\vec{\mathbf{P}}$ be a finite collection of multitiles. Let $\mu > 0$. Then after discarding tiles $\vec{P}$ such that $\langle f_4, \tilde{\phi}_P^3 \rangle = 0$, there exists a partition,*

$$\vec{\mathbf{P}} = \bigcup_{n : 2^n \leq \mathrm{size}_3((a_{P_3})_{\vec{P} \in \vec{\mathbf{P}}})} \bigcup_{T \in \mathcal{F}^{n,3}} T,$$

where $\mathcal{F}^{n,3}$ is a collection of trees such that $\mathrm{size}_3(T) \leq 2^{m+1}$ *and*

$$\sum_{T \in \mathcal{F}^{n,3}} |I_T| \lesssim |E_4| 2^{-2n} \left(2^{-n} |E_4|^{-1/2}\right)^{2/\mu}$$

Proof. See Corollary 6.4 in [8]. □

One gets nearly identical partition results for the P_2 and P_1 sequences using the energy results described for them, except that there is no presence of $(2^{-n})^{1/\mu}$ in these cases.

CHAPTER 3

RESTRICTED WEAK-TYPE INTERPOLATION

In this chapter, we discuss the so-called restricted weak-type interpolation theorems. They are valid for general n-linear operators, but we state them here for our specialized case.

Definition 3.0.6. *A tuple $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ is called admissible if*

1. $-\infty < \alpha_i < 1$ *for all* $i = 1, 2, 3, 4$
2. $\sum \alpha_i = 1$
3. *At most one* $\alpha_i < 0$.

We call an index i good if $\alpha_i \geq 0$ and bad if $\alpha_i < 0$. A good tuple is an admissible tuple without a bad index. A bad tuple is a tuple with a bad index.

Definition 3.0.7. *We define the term* ***majorant*** *as follows.*

1. *If α and β are good tuples and there exists a j_0 such that*

$$\alpha_j < \beta_j \text{ for all } j \neq j_0,$$

   then we say that β is a majorant of α with index j_0.

2. *If α or β are bad tuples, we assume that j_0 is their bad tuple (if they are both bad, this j_0 is the same for both). In this case, we say that β is a majorant of α with index j_0 if*

$$\alpha_j < \beta_j \text{ for all } j \neq j_0.$$

Definition 3.0.8. *Let E, E' be sets of finite measure. We say that E' is a major subset of E if $E' \subseteq E$ and $|E'| \geq \frac{1}{2}|E|$.*

Definition 3.0.9. *If E is a set of finite measure, we denote by $X(E)$ the space of functions supported on E such that $\|f\|_\infty \leq 1$.*

Definition 3.0.10. *If α is an admissible tuple, we say that a 4-linear form Λ is of restricted weak-type[1] α if for every sequence E_1, E_2, E_3, E_4 of subsets of $\mathbb{R}$ of finite measure, there exists a major subset E'_j of E_j for each bad index j (either 1 or 0) such that*

$$\Lambda(f_1, f_2, f_3, f_4) \lesssim |E'|^\alpha,$$

for all $f_i \in X(E_i)$, $i = 1, 2, 3, 4$, where we adopt the convention that $E'_i = E_i$ when i is a good index, and

$$|E'|^\alpha = |E'_1|^{\alpha_1}|E'_2|^{\alpha_2}|E'_3|^{\alpha_3}|E'_4|^{\alpha_4}.$$

Definition 3.0.11. *Suppose that a 4-linear form Λ is of restricted weak type α for some family of tuples $\alpha \in A$ which all have the same bad index j_0. Suppose further that the same major subset E'_{j_0} in the definition of restricted weak type can be used for all elements of A. Then we say that Λ is of uniformly restricted weak type.*

The basic idea here is that if $\Lambda(f_1, f_2, f_3, f_4) = \int T(f_1, f_2, f_3)f_4 dx$, then a good tuple can be written as $(1/p_1, 1/p_2, 1/p_3, 1/p_4)$ and corresponds to a standard Hölder type estimate for T, i.e. $L^{p_1} \times L^{p_2} \times L^{p_3} \to L^{p'_4}$. If a tuple had bad index 4, say, then the target space of T, $L^{p'_4}$, is necessarily not a Banach space since $1/p'_4 < 1$. Thus one cannot invoke immediately more standard interpolation results about mappings between Banach spaces. See, for example, [32] — an early paper of Cornell's own Bob Strichartz.[2]

[1] It is worth mentioning here that this is a slightly stronger definition of restricted weak type than others which appear in the literature, e.g. [13]. That said, there is a much stronger interpolation theorem available for this variant.

[2] It was quite a treat, years ago, to go looking for the original source of this result and to discover it was written by Bob!

The following theorem guarantees that one can interpolate multilinear restricted weak-type estimates as one can with usual multilinear estimates, provided the interpolated tuple is a good tuple.

Theorem 3.0.12. *Let $\alpha^{(1)}, ..., \alpha^{(4)}$ be admissible tuples, and let α be a* ***good*** *tuple such that*

$$\alpha = \theta_1 \alpha^{(1)} + ... + \theta_4 \alpha^{(4)},$$

where $0 < \theta_s < 1$ for $s = 1, 2, 3, 4$ and $\theta_1 + ... + \theta_4 = 1$. Suppose that Λ is of restricted weak type $\alpha^{(s)}$ for $s = 1, 2, 3, 4$. Then Λ is of restricted weak type α.

Proof. Consider the quantities

$$|\Lambda(f_1, f_2, f_3, f_4)|^{\theta_i} \lesssim \left(|E|^{\alpha^{(i)}}\right)^{\theta_i}$$

and multiply them together. □

The following theorem says that at good tuples on the interior of a convex, open set where a 4-linear form is of restricted weak type, then it is of strong type on the interior of the set.

Theorem 3.0.13. *Let $\alpha^{(1)}, ..., \alpha^{(4)}$ be tuples, and let α be a* ***good*** *tuple in the interior of the convex hull of $\alpha^{(1)}, ..., \alpha^{(4)}$. Suppose that Λ is of restricted weak-type $\alpha^{(s)}$ for $s = 1, 2, 3, 4$. Then Λ is of strong-type α.*

Proof. See [13, Corollary 1, pp 383–384]. □

These previous theorems actually hold for a weaker definition of restricted weak-type. They are not strong enough for our purposes because they require all the interpolated tuples to be good in order to produce estimates. Here are two replacements.

Lemma 3.0.14. *Suppose that a 4-linear form* Λ *is of uniformly restricted weak type* $\alpha^{(s)}$ *for* $s = 1, ..., 4$, *where all bad tuples, if they exist, have the same bad index. Suppose that*

$$\alpha = \theta_1 \alpha^{(1)} + ... + \theta_4 \alpha^{(4)},$$

where $0 < \theta_s < 1$ *for* $s = 1, 2, 3, 4$ *and* $\theta_1 + ... + \theta_4 = 1$. *Then* Λ *is of uniform restricted weak type for* $\{\alpha, \alpha^{(1)}, \alpha^{(2)}, \alpha^{(3)}, \alpha^{(4)}\}$. *Thus* Λ *is of uniform restricted weak type in the interior of the convex hull of the* $\alpha^{(s)}$.

Proof. Consider the quantities

$$|\Lambda(f_1, f_2, f_3, f_4)|^{\theta_i} \lesssim \left(|E|^{\alpha^{(i)}}\right)^{\theta_i}$$

and multiply them together, using the uniformity in the major subset. □

Lemma 3.0.15. *Suppose that* $\alpha^{(s)}$ *is a collection of tuples which are either good or bad with a fixed bad index for which* Λ *is of restricted weak type. Let*

$$\alpha := \theta_1 \alpha^{(1)} + ... + \theta_4 \alpha^{(4)},$$

where $0 < \theta_s < 1$ *for* $s = 1, 2, 3, 4$ *and* $\theta_1 + ... + \theta_4 = 1$. *We assume that some* $\alpha^{(j)}$ *is a majorant of* α *with index* j_0, *where*

1. *if* α *is good then* j_0 *an index for which* $\alpha_j > 0$.
2. *if* α *is bad then* j_0 *is that index.*

Then one has that Λ *is of restricted weak type* α *as well.*

Proof. See the appropriate appendix of [23]. It is also essentially [24, Lemma 3.10] □

These two lemmas give one the ability to interpolate between restricted weak-type estimates. However, we really want to be able to produce strong estimates for bad tuples. We prove the following lemma, which is just a special case of [24, Lemma 3.11].

Lemma 3.0.16. *Let α be a bad tuple with bad index 4. Suppose that our 4-linear form $\Lambda(f_1, f_2, f_3, f_4)$ satisfies a restricted weak-type estimate in an open neighborhood of α. Then if $\alpha_i = 1/p_i$ for $i = 1, 2, 3$ and $\alpha_4 = 1/p_4 = 1 - (1/p_4')$, we have*

$$\|T(f_2, f_2, f_3)\|_{p_4'} \leq C\|f_1\|_{p_1}\|f_2\|_{p_2}\|f_3\|_{p_3}$$

for all functions f_i supported on a set of finite measure.

Proof. The proof is [24, Lemma 3.11], but we include it here for completeness.

First, we may assume that none of the functions is in L^∞ for reasons as follows. Clearly, one can assume that $p_i \neq \infty$ for all $i \leq j$ and $p_i = \infty$ for all $i > j$ for some particular $j \in \{0, 1, 2, 3\}$. If $j \neq 3$, then $1/p_3 = 0$, then we may assume without loss of generality that $\|f_3\|_\infty = 1$. Then we may fix the function f_3 and apply the result when none of the $p_i = \infty$. So we assume no index is 0.

Now, we may assume without loss of generality that $\|f_i\|_{p_i} = 1$ for $i = 1, 2, 3$. Thus we need to show

$$\|T(f_2, f_2, f_3)\|_{p_4'} \lesssim 1.$$

We may assume that the functions are non-negative by linearity and the usual $f = f^+ - f^-$ trick, sacrificing a constant loss in the estimates. Moreover, modulo a measure-preserving transformation, we may assume that the f_i are non-increasing and supported on $(0, \infty)^3$. Since the f_i are supported on sets of finite measure,

[3] Confer results about non-increasing rearrangements, e.g. [3, Chapter 2].

Let

$$\chi_{k_i} := \chi_{(2^{k_i}, 2^{k_i+1}]}.$$

We may write our desired estimate as

$$\| \sum_{k_1,k_2,k_3} T(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3})\|_{p'_4} \lesssim 1.$$

Since $p'_4 \leq 1$, we have the inequality $|a+b|^{p'_4} \leq |a|^{p'_4} + |b|^{p'_4}$, and so it suffices to estimate

$$\sum_{k_1,k_2,k_3} \|T(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3})\|_{p'_4}^{p'_4} \lesssim 1.$$

Losing a factor of 3! in our estimates, we may estimate the sum over the region $k_1 \geq k_2 \geq k_3$. For a fixed k_1, k_2, k_3, let $\lambda > 0$ be arbitrary, and consider the set

$$E_4 = \{\Re T(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3}) > \lambda\}.$$

Let β be an admissible tuple near our α, close enough that it also has bad tuple n. Since Λ is of restricted weak-type β and $f_i\chi_{k_i} \in f_i(2^{k_i})X\left((2^{k_i-1}, 2^{k_i}]\right)$, we may find a major subset E'_4 of E_4 such that

$$|\Lambda(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3}, \chi_{E'_4})| \lesssim |E_4|^{\beta_4} \prod_{i=1}^{3} f_i(2^{k_i})2^{k_i\beta_i}.$$

By definition of E_4, we have

$$\lambda |E_4| \lesssim |E_4|^{\beta_4} \prod_{i=1}^{3} f_i(2^{k_i})2^{k_i\beta_i}$$

Solving for E_n and optimizing in β, one obtains

$$|E_4| \lesssim \lambda^{-p'_n} 2^{-\epsilon(k_1-k_3)} \min\left(\frac{F}{\lambda}, \frac{\lambda}{F}\right)^{\epsilon} \left(\prod_{i=1}^{3} f_i(2^{k_i})2^{k_i/p_i}\right)^{p'_4},$$

for some $\epsilon > 0$, where $F = \prod f_i(2^{k_i})$. By symmetry (and including the complex parts), this bound extends to

$$\{|T(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3})| > \lambda\}.$$

Now,

$$
\begin{aligned}
&\|T(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3})\|_{p_n'} \\
&= \left(\int_0^\infty p_n'\lambda^{p_n'-1} m(\{|T(f_1\chi_{k_1}, f_2\chi_{k_2}, f_3\chi_{k_3})| > \lambda\})d\lambda\right)^{1/p_4'} \\
&\lesssim \left(\int_0^\infty \lambda^{-1}2^{-\epsilon(k_1-k_3)} \min\left(\frac{F}{\lambda}, \frac{\lambda}{F}\right)^\epsilon\right)^{1/p_4'} \prod_{i=1}^3 f_i(2^{k_i})2^{k_i/p_i} \\
&= \left(\int_0^F \lambda^{-1+\epsilon}F^{-\epsilon}2^{-\epsilon(k_1-k_3)}\right)^{1/p_4'} \prod_{i=1}^3 f_i(2^{k_i})2^{k_i/p_i} \\
&+ \left(\int_F^\infty \lambda^{-1-\epsilon}F_\epsilon 2^{-\epsilon(k_1-k_3)}\right)^{1/p_4'} \prod_{i=1}^3 f_i(2^{k_i})2^{k_i/p_i} \\
&\lesssim 2^{-\epsilon(k_1-k_3)/p_n'} \prod_{i=1}^3 f_i(2^{k_i})2^{k_i/p_i},
\end{aligned}
$$

Thus we must show that

$$\sum_{k_1 \geq k_2 \geq k_3} 2^{-\epsilon(k_1-k_3)} \left(\prod_{i=1}^3 f_i(2^{k_i})2^{k_i/p_i}\right) \lesssim 1.$$

Write $s = k_1 - k_3$. For a fixed s there are at most $(1+s)^C$ choices for the k_i. Fix s and apply Hölder's inequality, observing that $p_1^{-1} + p_2^{-1} + p_3^{-1} = {p_4'}^{-1}$, to get

$$\sum_{s\geq 0}(1+s)^C 2^{-\epsilon s} \prod_{i=1}^3 \left(\sum_k (f_i(2^k))^{p_i}2^{k_i}\right)^{1/p_i}.$$

The sum in s is convergent, so it suffices to establish that

$$\prod_{i=1}^3 \left(\sum_k (f_i(2^k))^{p_i}2^k\right)^{1/p_i} \lesssim 1.$$

But this is done by

$$
\begin{aligned}
\prod_{i=1}^3 \left(2\sum_k (f_i(2^k))^{p_i}2^{k-1}\right)^{1/p_i} &\leq \prod_{i=1}^3 \left(2\sum_k \|f_i\chi_{k_i}\|_{p_i}^{p_i}\right)^{1/p_i} \\
&= \prod_{i=1}^3 2^{1/p_i}\|f_i\|_{p_i} \lesssim 1.
\end{aligned}
$$

$\square$

This lemma says that once one has a tiny open set worth of restricted weak-type estimates, one can get strong estimates on the interior for a class of functions like $C_c^\infty(\mathbb{R})$, which is enough to extend to T by the usual density arguments.

CHAPTER 4

MAIN RESULT FOR $M = 0$, $K_2 > K_1$

The application of the sizes, energies, and weak-type interpolation is fairly standard (for example, as in the article which inspired the present work, [8]), but we reproduce the procedure here.

We now state a basic lemma. It essentially comes from the intuition that one can estimate

$$\left|\sum_n a_n b_n c_n\right| \leq \|a_n\|_{\ell^2}\|b_n\|_{\ell^2}\|c_n\|_{\ell^\infty}$$

Lemma 4.0.17. *Suppose that T is a t-tree contained in $\vec{\mathbf{P}}$. This means it is a t'-lacunary tree for $t' \neq t$. As before, let*

1. $a_{P_1} = \langle f_3, \phi_P^1\rangle$
2. $a_{P_2} = \langle B_P(f_1, f_2), \phi_P^2\rangle$
3. $a_{P_3} = \left\langle f_4, \tilde{\phi}_P^3\right\rangle$

Then

$$\begin{aligned}
&\left|\sum_{\vec{P}\in T}\left|\frac{1}{|I_P|^{1/2}}\langle f_3, \phi_P^1\rangle\langle B_P(f_1, f_2), \phi_P^2\rangle\left\langle f_4, \tilde{\phi}_P^3\right\rangle\right|\right| \lesssim \\
&\sum_{\vec{P}\in T}\left|\frac{1}{|I_P|^{1/2}}\langle f_3, \phi_P^1\rangle\langle B_P(f_1, f_2), \phi_P^2\rangle\left\langle f_4, \tilde{\phi}_P^3\right\rangle\right| \lesssim \\
&\qquad |I_T|\text{size}_1((a_{P_1})_{\vec{P}\in T})\cdot\text{size}_2((a_{P_2})_{\vec{P}\in T})\cdot\text{size}_3((a_{P_3})_{\vec{P}\in T}).
\end{aligned}$$

Proof. By the definition of size,

$$\left(\sum_{\vec{P}\in T}|a_{P_{t'}}|^2\right)^{1/2} \lesssim |I_T|^{1/2}\text{size}_{t'}((a_{P_{t'}})_{\vec{P}\in T}),$$

for each $t' \neq t$. For t, one has that a single tile is a tree, and so

$$|a_{P_t}| \lesssim |I_P|^{1/2} \text{size}_t((a_{P_t})_{\vec{P} \in T}).$$

The claim then follows by the $\ell^2 \times \ell^2 \times \ell^\infty$ version of Hölder inequality (basically just Cauchy-Schwarz). $\square$

Supposing that $f_t \in X(E_t)$, this means it is enough (by restricted weak-type interpolation) to break up $\vec{\mathbf{P}}$ into trees T where one can produce the estimate

$$\sum_T |I_T| \text{size}_1((a_{P_1})_{\vec{P} \in T}) \cdot \text{size}_2((a_{P_2})_{\vec{P} \in T}) \cdot \text{size}_3((a_{P_3})_{\vec{P} \in T}) \lesssim |E_1|^{\alpha_1} |E_2|^{\alpha_2} |E_3|^{\alpha_3} |E_4|^{\alpha_4},$$

for an admissible tuple $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, where $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$.[1]

As per the restricted weak-type interpolation theorems, we are allowed to remove a certain subset from the E_n corresponding to a bad index (in the event that a bad index exists, or to any index in the event that no bad index exists). The indices 1 and 2 are to be handled differently from the indices 3 and 4: the functions f_1, f_2 are mixed together, and so their will have to be treated in a slightly different way than those for f_3, f_4. However, there is no difference between the methods used to handle 3 or 4.

4.1 Estimates when 3 or 4 is the bad index

We will describe in detail how to do this for index 4 being bad; the index 3 case can be done completely analogously.

[1]This last condition is clearly required since the operator in question behaves something like a pointwise product, and thus should satisfy Hölder-type estimates.

We will now define the exceptional set. For $C > 0$, define Ω_C as

$$\Omega_C := \bigcup_{i=1}^{4} \{x : M(1_{E_i}) \geq C|E_i|/|E_4|\},$$

where M is the usual Hardy–Littlewood maximal operator. For sufficiently large C, we can guarantee that $|E_4/\Omega_C| \geq \frac{1}{2}|E_4|$. Let E_4' be E_4/Ω_C for such a C.

Suppose that $f_1 \in X(E_1), f_2 \in X(E_2), f_3 \in X(E_3), f_4 \in X(E_4')$, and let $\vec{\mathbf{P}}$ be a finite rank 1 collection of tri-tiles. We partition $\vec{\mathbf{P}}$ as follows: let $\vec{\mathbf{P}}_l$ be the collection of tri-tiles $\vec{P}$ such that $I_{\vec{P}}$ satisfies

$$2^l \leq 1 + \frac{\text{dist}(I_{\vec{P}}, \mathbb{R}/\Omega)}{|I_{\vec{P}}|} \leq 2^{l+1}.$$

We will then have to sum over l. We shall find that we get an exponential gain of 2^{-l}, so this will not be an issue. Observe that, from our size estimates that for such collections of tiles,

$$\text{size}_1(a_{P_1})_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim \frac{|E_3|}{|E_4|}2^l$$

$$\text{size}_2(a_{P_2})_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim \frac{|E_1|^\theta |E_2|^{1-\theta}}{|E_4|}2^l$$

and

$$\text{size}_3(a_{P_3})_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim 2^{(1-M)l},$$

where M is the exponent from the definition of adaptedness to a tile.

Now, using Lemma 2.2.5 (and the appropriate analogues for P_1 and P_2), generate families $\mathcal{F}^{n,1}$, $\mathcal{F}^{n,2}$, and $\mathcal{F}^{n,3}$. After discarding tiles with $\langle f, \phi_P^t \rangle$, say, to zero, one can perform the partition,

$$\vec{\mathbf{P}}^l = \bigcup_{m_1,m_2,m_3} \mathbf{S}^{m_1} \cap \mathbf{S}^{m_2} \cap \mathbf{S}^{m_3},$$

where $\mathbf{S}^{m_t} := \bigcup_{T\in\mathcal{F}^{m,t}} T$ and we assume implicitly that

$$2^{m_t} \leq \text{size}_t((a_{P_t})_{\vec{P}\in\vec{\mathbf{P}}}).$$

One can further partition,

$$\vec{\mathbf{P}}^l = \bigcup_{j=1}^{3} \bigcup_{m_1,m_2,m_3 : m_j = \max\{m_1,m_2,m_3\}} \bigcup_{T \in \mathcal{F}^{m_j,j}} (T \cap \mathbf{S}^{m_1} \cap \mathbf{S}^{m_2} \cap \mathbf{S}^{m_3}).$$

Losing a factor of 3 in the estimates, we may drop the union over j and assume that

$$\vec{\mathbf{P}}^l = \bigcup_{m_1,m_2,m_3 : m_j = \max\{m_1,m_2,m_3\}} \bigcup_{T \in \mathcal{F}^{m_j,j}} (T \cap \mathbf{S}^{m_1} \cap \mathbf{S}^{m_2} \cap \mathbf{S}^{m_3}).$$

It is worth observing that $T \cap \mathbf{S}^{m_1} \cap \mathbf{S}^{m_2} \cap \mathbf{S}^{m_3}$ is still a tree with the same top as T and, by the sub-tree properties of the partition from Lemma 2.2.5, we have that its size is at most 2^{m_j+1}. Thus we must finally verify that

$$\sum_{m_1,m_2,m_3 : m_j = \max\{m_1,m_2,m_3\}} \sum_{T \in \mathcal{F}^{m_j,j}} |I_T| 2^{m_1+m_2+m_3} \lesssim 2^{-l} |E_1|^{\alpha_1} |E_2|^{\alpha_2} |E_3|^{\alpha_3} |E_4|^{\alpha_4}.$$

Suppose that $a_1 + a_2 + a_3 = 1 = \frac{1-a_1}{2} + \frac{1-a_2}{2} + \frac{1-a_3}{2}$ for $0 \leq a_1, a_2, a_3 \leq 1$. Let $0 < \theta < 1$. By Lemma 2.2.5 and its two variants, we have that

$$\sum_{T \in \mathcal{F}^{m_j,j}} |I_T| \lesssim 2^{-2m_j} \left(|E_3|\right)^{\frac{1-a_1}{2}} \left(|E_1|^{\theta} |E_2|^{1-\theta}\right)^{\frac{1-a_2}{2}} \tag{4.1.1}$$

$$\times \left(|E_4| \left(2^{-m_j} |E_4|^{-1/2}\right)^{2/\mu}\right)^{\frac{1-a_3}{2}} \tag{4.1.2}$$

Since we assumed implicitly that

$$2^{m_t} \leq \mathrm{size}_t((a_{P_t})_{\vec{P} \in \vec{\mathbf{P}}}),$$

we have that, for the same a_1, a_2, a_3:

$$\begin{aligned} 2^{m_1+m_2+m_3} &= 2^{m_1(1-a_1)+m_2(1-a_2)+m_3(1-a_3)} 2^{m_1 a_1 + m_2 a_2 + m_3 a_3} \\ &\leq 2^{m_j} \prod_{i \neq j} \mathrm{size}_i((a_{P_i})_{\vec{P} \in \vec{\mathbf{P}}})^{a_i} 2^{m_i(1-a_i)} \end{aligned}$$

Thus by summing up the geometric sums over m_i, which cap out at m_j, one has

$$\sum_{m_1,m_2,m_3 : m_j = \max\{m_1,m_2,m_3\}} \sum_{T \in \mathcal{F}^{m_j,j}} |I_T| 2^{m_1+m_2+m_3} \lesssim$$

$$\prod_{i \neq j} \mathrm{size}_i((a_{P_i})_{\vec{P} \in \vec{\mathbf{P}}})^{a_i} \sum_{m_j} 2^{m_j} \left(\prod_{i \neq j} 2^{m_j(1-a_i)}\right) \sum_{T \in \mathcal{F}^{m_j,j}} |I_T|$$

Now, plugging in (4.1.1), and summing over the final geometric series and carefully doing some arithmetic on the exponents, one can majorize the previous expression by

$$\left(\prod_{i=1}^{3} \mathrm{size}_i((a_{P_i})_{\vec{P}\in\vec{\mathbf{P}}})^{a_i}\right) \left(\mathrm{size}_j((a_{P_j})_{\vec{P}\in\vec{\mathbf{P}}})\right)^{-(1-a_3)/\mu}$$
$$\times |E_3|^{(1-a_1)/2}|E_1|^{\theta(1-a_2)/2}|E_2|^{(1-\theta)(1-a_2)/2}|E_4|^{(1-a_3)/2}|E_4|^{-(1-a_3)/\mu}|E_4|^{-1}.$$

Let $(1-a_3)/\mu = \epsilon$. Observe that the presence of the $\mathrm{size}_j^{-\epsilon}$ term is harmless except that it effectively changes the factor of $|E_4|^{-\epsilon}$ to $|E_3|^{-\epsilon}$ if $j = 1$, to $|E_1|^{-\theta\epsilon}|E_2|^{-(1-\theta)\epsilon}$ if $j = 2$ or leaves this factor if $j = 3$. This can be remedied quite easily. Supposing that $j = 1$, pick $\alpha_1' = \alpha_1 + \epsilon$ (which is ok for "most" choices of α_1 since μ can be taken very large) and making the appropriate change $\alpha_3' = \alpha_1 - \epsilon$. Thus the $-\epsilon$ can always be pushed onto E_4. But the key point is that one gets a weak-type estimate for all ϵ, so one can get estimates arbitrarily close to $\epsilon = 0$. Thus we ignore this technicality. We can thus majorize the previous expression by quantities arbitrarily close to

$$|E_3|^{(1+a_1)/2}|E_1|^{\theta(1+a_2)/2}|E_2|^{(1-\theta)(1+a_2)/2}|E_4|^{(1+a_3)/2}|E_4|^{-1}$$

whenever $0 < a_1, a_2, a_3 < 1$ with $a_1 + a_2 + a_3 = 1$ and $0 < \theta < 1$. All the associated tuples are admissible tuples, and hence our 4-linear form Λ is of restricted weak-type for all such α, θ pairs. If one picks:

1. $a_1 = 2\alpha_3 - 1$,
2. $a_3 = 2\alpha_4 - 1$,
3. $a_2 = 2(\alpha_1 + \alpha_2) - 1$, and
4. $\theta = \alpha_1/(\alpha_1 + \alpha_2)$,

then the previous estimate becomes

$$|E_1|^{\alpha_1}|E_2|^{\alpha_2}|E_3|^{\alpha_3}|E_4|^{\alpha_4'},$$

where $\alpha_4' = \alpha_4 - 1$. Of course the sum of the exponents is then 1, and hence our 4-linear form is of restricted weak-type α whenever $0 < \alpha_3 < 1$, $-1 < \alpha_4' < 0$, $0 \leq \alpha_1, \alpha_2 \leq 1$, $0 < \alpha_1 + \alpha_2 < 1$, and $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4' = 1$. In particular, one gets a restricted weak-type estimate for 4-tuples arbitrarily close to

$$(0, 1, \tfrac{1}{2}, -\tfrac{1}{2}) \quad (0, \tfrac{1}{2}, 1, -\tfrac{1}{2}) \quad (1, 0, \tfrac{1}{2}, -\tfrac{1}{2}) \quad (\tfrac{1}{2}, 0, 1, -\tfrac{1}{2}) \quad (1, \tfrac{1}{2}, 0, -\tfrac{1}{2}) \quad (\tfrac{1}{2}, 1, 0, -\tfrac{1}{2}).$$

One can do precisely the same analysis for 3 being the bad index to get restricted weak-type estimates

$$(0, 1, -\tfrac{1}{2}, \tfrac{1}{2}) \quad (0, \tfrac{1}{2}, -\tfrac{1}{2}, 1) \quad (1, 0, -\tfrac{1}{2}, \tfrac{1}{2}) \quad (\tfrac{1}{2}, 0, -\tfrac{1}{2}, 1) \quad (1, \tfrac{1}{2}, -\tfrac{1}{2}, 0) \quad (\tfrac{1}{2}, 1, -\tfrac{1}{2}, 0).$$

4.2 Estimates when 1 or 2 is the bad index

Now, the operator can be estimated in nearly the same way, although there are some minor changes which we now describe. We prove the estimates for 2 being the bad index. The case for 1 being the bad index is completely analogous.

We define the exceptional set

$$\Omega_C = \bigcup_{j=1}^{4} \{M(\chi_{E_j}) > C|E_j|/|E_2|\},$$

where again M is the Hardy–Littlewood maximal operator. For sufficiently large C, we can define $E_2' = E_2/\Omega_C$ to get an appropriate major subset.

Now, we make two assumptions of a similar type to the ones we made before: we restrict to tiles $\vec{P}$ with

$$2^k \leq 1 + \frac{\operatorname{dist}(I_{\vec{P}}, \mathbb{R}/\Omega_C)}{|I_{\vec{P}}|} \leq 2^{k+1}$$

and to tiles $\vec{Q}$ with

$$2^{k'} \leq 1 + \frac{\mathrm{dist}(I_{\vec{Q}}, \mathbb{R}/\Omega_C)}{|I_{\vec{Q}}|} \leq 2^{k'+1},$$

which is harmless provided we get summability in k, k'. One then proceeds in exactly the same fashion, except that one needs to make the following changes to the size_3 and energy_3:

$$\mathrm{size}_3((a_{P_2})_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim 2^{(-M\theta)k},$$

where one must use the crude estimate $\int_{E_j} \tilde{\chi}^M_{I_{\vec{Q}}} \leq |I_{\vec{Q}}|$. We are already choosing M depending on the exponent parameters, so the presence of θ is ok, provided it is nonzero. We also get

$$\mathrm{energy}_3((a_{P_2})_{\vec{P}\in\vec{\mathbf{P}}}) \lesssim 2^{-M\theta k'}|E_1|^{(2-\theta)/2}|E_2|^{(\theta-1)/2},$$

for some $0 < \theta < 1$. One gets summability in k, k', so we may ignore their presence. The estimate one gets as before (ignoring the small factor $1/\mu$) is

$$\frac{|E_3|^{(1+a_1)/2}|E_4|^{(1+a_2)/2}}{|E_2|^{1-a_3}}\left(|E_1|^{(2-\theta)/2}|E_2|^{(\theta-1)/2}\right)^{1-a_3}.$$

Now pick

1. $a_1 = 2\alpha_3 - 1$,
2. $a_2 = 2\alpha_4 - 1$,
3. $a_3 = 2(\alpha_1 + \alpha_2) + 1$, and
4. $\theta = (3\alpha_1 + 2\alpha_2)/(\alpha_1 + \alpha_2)$.

This numerology transforms the previous line to

$$|E_3|^{\alpha_3}|E_4|^{\alpha_4}|E_1|^{\alpha_1}|E_2|^{\alpha_2}.$$

i.e. producing a weak-type α estimate. One may now check that tuples arbitrarily close to the following are available:

$$(\tfrac{1}{2}, -\tfrac{3}{2}, 1, 1) \quad (1, -\tfrac{3}{2}, \tfrac{1}{2}, 1) \quad (1, -\tfrac{3}{2}, 1, \tfrac{1}{2}).$$

By doing the same analysis for 1 being the bad index, one gets

$$(-\tfrac{3}{2}, \tfrac{1}{2}, 1, 1) \quad (-\tfrac{3}{2}, 1, \tfrac{1}{2}, 1) \quad (-\tfrac{3}{2}, 1, 1, \tfrac{1}{2}).$$

4.3 Main Result

By the work of the previous two sections and the weak-type interpolation theorems discussed at the beginning of this chapter, we have proven the following theorem.

Theorem 4.3.1. *Define $T(f_1, f_2, f_3)$ by*

$$\Bigg| \sum_{i \gg j} \int \int_{\mathbb{R}^3} \hat{f}_1(\xi_1)\hat{f}_2(\xi_2)\hat{f}_3(\xi_3)\theta_i(\xi_2 - \xi_1) \\ \phi_j\left(\xi_3 - \frac{\xi_1 + \xi_2}{2}\right) e^{2\pi i x(\xi_1+\xi_2+\xi_3)} d\xi 1_{j \geq N_2(x)} \Bigg|,$$

where $N_2(x)$ is an arbitrary, integer-valued function on $\mathbb{R}$. Let D denote the interior of the convex hull in $\{(\alpha_1, \alpha_2, \alpha_3, \alpha_4) : \sum \alpha_i = 1\}$ of the 4-tuples given at the end of the previous two sections. Suppose that $1 < p_1, p_2, p_3 < \infty$ and $1/p_4' = 1 - 1/p_4 = \sum_i^3 1/p_i$ are such that $(1/p_1, 1/p_2, 1/p_3, 1/p_4)$ is in D. Then

$$T : L^{p_1} \times L^{p_2} \times L^{p_2} \rightarrow L^{p_4'}.$$

Corollary 4.3.2. *Suppose that T is as in the previous theorem. Then $T : L^2 \times L^2 \times L^2 \rightarrow L^{2/3}$ is bounded.*

This corollary is of particular interest since we get a strong bound into $L^{2/3}$. All the "trivial" methods of estimation require putting one of the f_i into L^∞ and then using previous methods to make estimations on the remaining objects; however, the only estimates available have either the other f_i in L^p and L^q where either $p^{-1} + q^{-1} = 1$ (if $f_2 \in L^\infty$ and applying Hölder on the tensor product of two maximal operators) or $p^{-1} + q^{-1} > 3/2$ (either f_1 or f_3 in L^∞ and applying time-frequency analysis in the spirit of of Lacey's original paper on the maximal bilinear operator, [17], or the relevant special case of Demeter, Tao, Thiele, [8]). Either way, one cannot produce estimates using the prior estimates so that the target space is actually being $L^{2/3}$.

CHAPTER 5

SIZES FOR $M = 0$, $K_2 \leq K_1$

5.1 Estimate on a Single Tree

We denote by $X(E)$ the space of all smooth functions supported in E and bounded by 1. The idea is to think of functions in $X(E)$ as smooth "characteristic functions" of E. After dualizing, our four-linear form (for the $m = 0$ Taylor term) with inputs f_1,f_2,f_3, and f_4 (which are in $X(E_j)$, respectively) is given by

$$\sum_{Q \in \mathbf{Q}} \frac{1}{|I_Q|^{1/2}} \left\langle f_1, \phi_Q^1 \right\rangle \left\langle f_2, \phi_Q^2 \right\rangle \left\langle M_Q(f_3) b_{P,Q} f_4, \phi_Q^3 1_{\{x:|I_Q|>2^{N_1(x)}\}} \right\rangle,$$

where

$$M_Q(f_3) = \sum_{P \in \mathbf{P}: \omega_{P_2} \supsetneq \omega_{Q_3}} \left\langle f_3, \phi_P^1 \right\rangle \phi_P^1 1_{\{x:|I_P|>2^{N_2(x)}\}},$$

$$b_{P,Q} = e^{\frac{|I_P|}{|I_Q|} \pi i n_2 \ell_{Q_2}},$$

and $\mathbf{P}$ and $\mathbf{Q}$ are some finite families of multitiles and $\mathbf{P}$ has rank 1. For ease of writing, we will use the following notation:

$$\tilde{\phi}_Q^3 := \phi_Q^3 1_{\{x:|I_P|>2^{N_1(x)}\}},$$

$$\tilde{\phi}_P^1 := \phi_P^1 1_{\{x:|I_P|>2^{N_2(x)}\}}.$$

Now, we saw in Lemma 4.0.17 that the key ingredient is to restrict the sum of our model to a sum over a tree. Thus we do just that. So, let $\mathbf{T}$ be a 1-, 2-, or 3-tree and consider

$$\sum_{Q \in \mathbf{T}} \frac{1}{|I_Q|^{1/2}} \left\langle f_1, \phi_Q^1 \right\rangle \left\langle f_2, \phi_Q^2 \right\rangle \left\langle M_Q(f_3) b_{P,Q} f_4, \tilde{\phi}_Q^3 \right\rangle.$$

Our first move is to simplify $b_{P,Q}$.

Suppose that $\mathbf{T}$ is either a 1-tree or a 2-tree, i.e. it is a lacunary 3-tree.

Case I: Suppose that $\mathbf{T}$ is a 2-tree. Then all the ω_{Q_2} intervals "morally" hit ω_T, the frequency interval for the top of the tree. Thus

$$\frac{\ell_{Q_2}}{|I_Q|} \approx K_T = \frac{\ell_T}{|I_T|},$$

where K_T is of course some integer which depends on T — modulo a constant loss, we may assume that K_T is the same for all Q. The coefficient $b_{P,Q}$ then becomes

$$b_{P,Q} = e^{\pi i n_2 |I_P| K_T}.$$

But now one can attach this coefficient to ϕ_P^1. This has no effect on the Fourier support or any of the L^p norms or the constants in the definition of adaptedness. Hence it is completely harmless and can be ignored, which we shall do.

Case II: Suppose that $\mathbf{T}$ is a 1-tree. Then all the ω_{Q_1} intervals "morally" hit ω_T, the frequency interval for the top. But then this means that $\mathbf{T}$ is 2-lacunary, i.e. $\mathrm{dist}(10\omega_T, 10I_Q|) \approx |I_Q|^{-1}$. Hence

$$\frac{\ell_{Q_2}}{|I_Q|} \approx K_T + \frac{b}{|I_Q|}$$

for some b. There are only finitely many possibilities for b, and so, modulo a finite loss in the estimate, we may assume that all the Q in $\mathbf{T}$ have the same b. Hence

$$b_{P,Q} = e^{\pi i n_2 |I_P| K_T} e^{b\pi i n_2 \frac{|I_P|}{|I_Q|}}.$$

The first factor here can be dealt with as in the previous case. The second factor can be split again as a Taylor series. We label the terms in this expansion by m'. For the m'-term, one gets a factor of $|I_P|/|I_Q| = 2^{-m'k}$ for some $k > 10$ as well as a factor of $n_2^{m'}$. The polynomial in n_2 is harmless, provided we only need finitely many m'. For $m' \gg 0$, one again defines $N_k(x)$ and attach it to ϕ_P^1, as we did in

Section 1.3.2. Then, after performing a discretization again with respect to ϕ_Q^3, one gets that our current sizes will be the same as the sizes for the model we derived in Section 1.3.2, only with the factor $b_{P,Q}$ replaced with $2^{-m'k}$. Thus the Taylor remainder (in m') can be handled in precisely the same way. The intermediate m', however, will require more robust techniques, as we have stated several times. We discuss for the rest of this section the $m' = 0$ case.

5.2 3-size, $m' = 0$

We are interested in estimating the j-size of a tree, for $j = 1, 2, 3$; the 1- and 2-sizes are standard and are defined precisely as in the earlier sections; the estimates one produces are exactly as in Lemma 2.1.4. Our 3-size is analogous: given a family of tiles $\mathbf{S}$, we define

$$\mathrm{size}_3(\mathbf{S}) := \sup_{\mathbf{T} \subset \mathbf{S}} \frac{1}{|I_T|^{1/2}} \left(\sum_{Q \in \mathbf{T}} \left| \langle M_Q(f_3) \tilde{\phi}_Q^3, f_4 \rangle \right|^2 \right)^{1/2},$$

where the supremum is taken over Q-trees in $\mathbf{S}$ which are $1-$ or $2-$trees, i.e. they are 3-lacunary. Recall that $b_{P,Q}$ does not appear as a result of the previous section's discussion. We will additionally require that all trees either "grow up" or "grow down"; specifically, we require that a tree which is lacunary in Q_3 to have frequency intervals lying either all above ω_T or all below. By the triangle inequality, this results in a constant loss in our estimates.

So, suppose that $\mathbf{T}$ is a tree which is lacunary in the third position which grows up. We specifically want to obtain appropriate estimates for

$$\frac{1}{|I_T|^{1/2}} \left(\sum_{Q \in \mathbf{T}} \left| \left\langle M_Q(f_3) \tilde{\phi}_Q^3, f_4 \right\rangle \right|^2 \right)^{1/2}.$$

The main estimate we will show is that when the 3-size of a collection of Q-tiles is roughly 2^s for some s, then the 3-size is controlled by a product of p- and q-averages of the two functions over the top a particular sub-tree where $p^{-1}+q^{-1}=1-\epsilon$, i.e. with the target just above L^1. The primary obstacle to estimating the above quantity is that the function $M_Q(f_3)$ depends on the tile Q. With this in mind, we make the following definition:

$$\mathcal{P}_{\mathbf{T}} = \left\{ P \in \mathbf{P} : \omega_{P_2} \supsetneq \omega_{Q'_3} \text{ for some } Q' \in \mathbf{T} \right\}.$$

The family $\mathcal{P}_{\mathbf{T}}$ is the collection of all P-tiles which could possibly be involved in the above sum. We will show that if we replace the sum over all $P \in \mathbf{P}$ such that $\omega_{Q_3} \subseteq \omega_{P_2}$ with the sum over $\mathcal{P}_{\mathbf{T}}$, then the size is controlled by particular averages of f_3 and f_4. More precisely, we prove the following proposition.

Proposition 5.2.1. *Let f_3, f_4 be functions in $X(E_3)$ and $X(E_4)$ respectively, with $X(E)$ as defined at the beginning of this section. Suppose that $\tilde{\mathbf{T}}$ is a tree of Q-multitiles which is lacunary in the third position. Suppose further that $s \in \mathbb{Z}$ and*

$$2^s \leq \mathrm{size}_3(\tilde{\mathbf{T}}) \leq 2^{s+1}.$$

Then for any $\epsilon \in (0,1)$ and $p, q \in (1,\infty)$ with $1/p + 1/q = 1-\epsilon$,

$$\mathrm{size}_3(\tilde{\mathbf{T}}) \lesssim \frac{\|f_3 \tilde{\chi}_{I_T}\|_p \|f_4 \tilde{\chi}_{I_T}\|_q}{|I_T|^{1/p+1/q}},$$

where $\mathbf{T}$ is the maximizing subtree for $\mathrm{size}_3(\tilde{\mathbf{T}})$ (which exists since $\tilde{\mathbf{T}}$ is a finite collection of multitiles). Here $\tilde{\chi}_{I_T}$ is a positive function adapted to I_T, i.e. which decays rapidly away from I_T in units of $|I_T|$.

Proof. First, we suppose that $\mathbf{T}$ is the tree which is the maximizer of $\mathrm{size}_3(\tilde{\mathbf{T}})$. We begin by dualizing the sum over $Q \in \mathbf{T}$; in particular, it suffices to estimate

$$\left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \sum_{P \in \mathcal{P}_{\mathbf{T}}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \tilde{\phi}_Q^3, f_4 \right\rangle \right|,$$

where a_Q is an arbitrary sequence in ℓ^2 with ℓ^2-norm less than 1 which maximizes this quantity. In particular, we know that

$$|a_Q| = \frac{\left|\left\langle \sum_{P \in \mathcal{P}_{\mathbf{T}}} \langle f_3, \phi_P^1 \rangle \tilde{\phi}_P^1 \tilde{\phi}_Q^3, f_4 \right\rangle\right|}{\#}, \tag{5.2.1}$$

where $2^s |I_T|^{1/2} \le \# \le 2^{s+1}|I_T|^{1/2}$ by the requirement that $\mathrm{size}_3(\mathbf{T}) = \mathrm{size}_3(\tilde{\mathbf{T}}) \approx 2^s$.

For any integer n, let $I_{T,n}$ denote the interval I_T translated by $n|I_T|$ units. We begin by picking any m, n and assuming that $f_3 = f_3 \chi_{I_{T,m}}$ and $f_4 = f_4 \chi_{I_{T,n}}$. We expect (and indeed will show below) substantial decay in $|n|$ and $|m-n|$ when $|m-n| \gg 0$ and when $|n| \gg 0$. Indeed, $m \approx n$ and $n \approx 0$ will be the main terms, while $|m-n| \gg 0$, $n \gg 0$ can be estimated using crude estimates. For sake of completeness, we include all possible cases.

For ease of writing, let $F_3 = \sum_{P \in \mathcal{P}_{\mathbf{T}}} \langle f_3, \phi_P^1 \rangle \tilde{\phi}_P^1$. Then observe that

$$\begin{aligned} \left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \tilde{\phi}_Q^3, F_3 f_4 \right\rangle \right| &= \left| \frac{1}{|I_T|^{1/2}} \left\langle \sum_{Q \in \mathbf{T}} a_Q \tilde{\phi}_Q^3, F_3 f_4 \right\rangle \right| \\ &\lesssim \frac{1}{|I_T|^{1/2}} \int \left| \sum_{Q \in \mathbf{T}} a_Q \tilde{\phi}_Q^3(x) \right| |F_3(x)||f_4(x)| dx. \end{aligned}$$

So for $p, q, r > 1$ with $p^{-1} + q^{-1} + r^{-1} = 1$, Hölder's inequality gives us

$$\frac{1}{|I_T|^{1/2}} \int \left| \sum_{Q \in \mathbf{T}} a_Q \tilde{\phi}_Q^3 \right| |F_3(x)||f_4(x)| dx \tag{5.2.2}$$

$$\le \frac{1}{|I_T|^{1/2}} \|F_3\|_{L^p(I_{T,n})} \|f_4\|_{L^q(I_{T,n})} \left\| \sum_{Q \in \mathbf{T}} a_Q \tilde{\phi}_Q^3 \right\|_{L^r(I_{T,n})}, \tag{5.2.3}$$

where the factor of $\chi_{I_{T,n}}$ comes from our assumption about f_4 at the beginning of this proof.

Third term of (5.2.3), when $|n| \ge 2$**.** We begin by estimating the last term in (5.2.3). If $|n| \ge 2$, then by invoking the extreme decay of the functions $\tilde{\phi}_Q^3$

(obviously, $|\tilde{\phi}^3_Q| \leq |\phi^3_Q|$) away from I_T we may put absolute values inside and make the following coarse pointwise estimate: for some large integer N (by splitting scale-by-scale),

$$\left|\sum_{Q\in\mathbf{T}} a_Q \tilde{\phi}^3_Q\right| \chi_{I_{T,n}} \lesssim \frac{1}{|I_T|^{1/2}} \sum_{Q\in\mathbf{T}} |a_Q| \frac{|I_T|^{1/2}}{|I_Q|^{1/2}} \tilde{\chi}^N_{I_Q} \chi_{I_{T,n}}$$
$$\leq \frac{1}{|I_T|^{1/2}} \sum_{Q\in\mathbf{T}} |a_Q| \frac{1}{n^N} \left(\frac{|I_Q|}{|I_T|}\right)^{N-1/2}.$$

Applying Cauchy–Schwarz and using the fact that $(a_Q)_Q$ has ℓ^2-norm majorized by 1, one gets that

$$\left\|\sum_{Q\in\mathbf{T}} a_Q \tilde{\phi}^3_Q\right\|_{L^r(I_{T,n})} \lesssim \frac{1}{n^N} \frac{1}{|I_T|^{\frac{1}{2}-\frac{1}{r}}},$$

which produces heavy decay in $|n|$ as we anticipated.

Third term of (5.2.3), when $|n| \leq 2$. Now, supposing that $|n| \leq 2$, we replace $L^r(I_{T,n})$ with simply L^r, losing at most a factor of 5 in our estimates. First, we observe the pointwise estimate,

$$\left|\sum_{Q\in\mathbf{T}} a_Q \tilde{\phi}^3_Q\right| \leq \sup_{k\in\mathbb{Z}} \left|\sum_{Q\in\mathbf{T}, |I_Q|>2^k} a_Q \phi^3_Q\right|.$$

Since the Q_3 tiles are organized in a lacunary tree, the summation in this last expression can be written as the sum over all $Q \in \mathbf{T}$ convolved with a function whose Fourier support is in some interval of the form $[-2^k, 2^k]$, modulo a fixed constant. This can then be majorized by $M\left(\sum_{Q\in\mathbf{T}} a_Q \phi^3_Q\right)$, where M is the usual Hardy–Littlewood maximal operator. Thus we have that, for $r > 1$,

$$\left\|\sum_{Q\in\mathbf{T}} a_Q \tilde{\phi}^3_Q\right\|_{L^r} \leq \left\|M\left(\sum_{Q\in\mathbf{T}} a_Q \phi^3_Q\right)\right\|_{L^r} \lesssim \left\|\sum_{Q\in\mathbf{T}} a_Q \phi^3_Q\right\|_{L^r}.$$

We now apply a nice trick. Dualize with a function $g \in L^{r'}$. Then

$$\left|\int \sum_{Q\in\mathbf{T}} a_Q \phi^3_Q(x) g(x) dx\right| = \left|\sum_{Q\in\mathbf{T}} a_Q \langle g, \phi^3_Q\rangle\right| = \left|\int \sum_{Q\in\mathbf{T}} \frac{a_Q}{|I_Q|^{1/2}} \frac{\langle g, \phi^3_Q\rangle}{|I_Q|^{1/2}} 1_{I_Q}(x) dx\right|.$$

Applying the Cauchy–Schwarz inequality (pointwise for each x), we may bound the above by

$$\int \left(\sum_{Q \in \mathbf{T}} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \left(\sum_{Q \in \mathbf{T}} \frac{|\langle g, \phi_Q^3 \rangle|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} dx.$$

Invoking Hölder's inequality (with respect to the integral), we may majorize this by

$$\left\| \left(\sum_{Q \in \mathbf{T}} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_r \left\| \left(\sum_{Q \in \mathbf{T}} \frac{|\langle g, \phi_Q^3 \rangle|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_{r'} .$$

The second term here is a Littlewood–Paley-type square function of g, which is bounded from $L^s \to L^s$ for all $s > 1$ (see, for example, [22]); since $\|g\|_{r'} \leq 1$, we may clearly estimate just the first term in this expression. Now, if we let $\mathcal{I}$ denote the collection of all I_Q such that $Q \in \mathbf{T}$, then (after observing that each $I \in \mathcal{I}$ corresponds to a unique tile Q)

$$\left\| \left(\sum_{Q \in \mathbf{T}} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_r \leq |I_T|^{1/r} \sup_{I_0 \in \mathcal{I}} \frac{1}{|I_0|^{1/r}} \left\| \left(\sum_{I_Q \in \mathcal{I} : I \subseteq I_0} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_r$$

This last expression is a BMO-norm, and so by a discrete variant of the John–Nirenberg inequality, e.g. as shown in [22], we may replace r with 2. The last term then becomes

$$|I_T|^{1/r} \sup_{I_0 \in \mathcal{I}} \frac{1}{|I_0|^{1/2}} \left(\sum_{I_Q \in \mathcal{I} : I \subseteq I_0} |a_Q|^2 \right)^{1/2} .$$

Invoking (5.2.1) yields

$$|I_T|^{1/r} \frac{1}{\#} \sup_{I_0 \in \mathcal{I}} \frac{1}{|I_0|^{1/2}} \left(\sum_{I_Q \in \mathcal{I} : I \subseteq I_0} \left| \left\langle \sum_{P \in \mathcal{P}_{\mathbf{T}}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \tilde{\phi}_Q^3, f_4 \right\rangle \right|^2 \right)^{1/2} .$$

The quantity involving the supremum is now clearly a supremum over sub-trees and hence is controlled by the $\text{size}_3(\mathbf{T})$ which is, in turn, controlled by 2^{s+1}. Thus we may majorize the third term in (5.2.3) by

$$|I_T|^{1/r} \frac{2^{s+1}}{\#} \leq |I_T|^{1/r} \frac{2^{s+1}}{2^s |I_T|^{1/2}} \lesssim \frac{1}{|I_T|^{1/2 - 1/r}}.$$

Thus the third term in (5.2.3) is easily majorized by

$$\min\left\{1, \frac{1}{n^{N-1/2}}\right\} \frac{1}{|I_T|^{1/2-1/r}}.$$

First term of (5.2.3), when $|m-n| \geq 5$. We now estimate the first term in (5.2.3). Suppose first that $|m-n| \geq 5$. We wish to estimate

$$\left\| \sum_{P \in \mathcal{P}_{\mathbf{T}}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \chi_{I_{T,n}} \right\|_p .$$

We split the sum into two those parts where $I_P \subseteq I_{T,j}$, for some integer j. By the triangle inequality, it suffices to estimate each piece separately. We begin with the assumption that $|j-n| \geq 2$. First, dualize with a function $g \in L^{p'}$. Then, as before,

$$\left| \sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \left\langle f_3, \phi_P^1 \right\rangle \langle g \chi_{I_{T,n}}, \tilde{\phi}_P^1 \rangle \right| \tag{5.2.4}$$

$$\lesssim \left| \int \sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{\langle f_3, \phi_P^1 \rangle}{|I_P|^{1/2}} 1_P(x) \frac{\langle g \chi_{I_{T,n}}, \tilde{\phi}_P^1 \rangle}{|I_P|^{1/2}} 1_{I_P}(x) \right| \tag{5.2.5}$$

$$\leq \left\| \left(\sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{| \left\langle f_3, \phi_P^1 \right\rangle |^2}{|I_P|} 1_{I_P} \right)^{1/2} \right\|_p \left\| \left(\sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{|\langle g \chi_{I_{T,n}}, \tilde{\phi}_P^1 \rangle|^2}{|I_P|} 1_{I_P} \right)^{1/2} \right\|_{p'} \tag{5.2.6}$$

We observe that, if $|j-m| \geq 2$, then we may coarsely estimate the first factor in

(5.2.6).

$$\begin{aligned}
&\left\| \left(\sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{|\langle f_3, \phi_P^1 \rangle|^2}{|I_P|} 1_{I_P} \right)^{1/2} \right\|_p \\
&\lesssim \sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{\|f_3\|_p \|\phi_P^1 \chi_{I_{T.m}}\|_{p'}}{|I_P|^{1/2}} 1_{I_P} \\
&\lesssim \|f_3\|_p \sum_{k=0}^{\infty} \left\| \sum_{I_P \subseteq I_{T,j} : |I_P| = 2^{-k}|I_T|} \left(\frac{|I_P|}{|j-m||I_T|} \right)^N \frac{|I_T|^{1/p'}}{|I_P|} 1_{I_P} \right\|_p \\
&\lesssim \|f_3\|_p \sum_{k=0}^{\infty} \sum_{I_P \subseteq I_{T,j} : |I_P| = 2^{-k}|I_T|} 2^{-kN} \frac{2^{k/p'} |I_P|^{1/p'}}{|I_P||j-m|^N} \|1_{I_P}\|_p.
\end{aligned}$$

Now, we must mention briefly that the P_2 tiles coming from $\mathcal{P}_{\mathbf{T}}$ have frequency intervals which contain some ω_{Q_3}. By virtue of the scale separation and the fact that the Q_3 tiles form a lacunary tree, it follows that restricting to the family of P_2 tiles whose time intervals are contained in I_T all intersect the top of the Q_3 tree. Similarly, the P_2 tiles whose time intervals are contained in $I_{T,j}$ all intersect the top of the Q_3 tree translated by $j|I_T|$. Thus the P_1 tiles restricted to having time intervals in $I_{T,j}$ actually form a lacunary tree as well. Hence in the previous expression, each I_P corresponds to a unique tile in $\mathcal{P}_{\mathbf{T}}$, and so in particular there are 2^k possible I_P's sitting inside I_T. So, the last expression above is majorized by

$$\|f_3\|_p \sum_{k=0}^{\infty} 2^{-k(N-1-1/p')} \frac{|I_P|^{1/p+1/p'}}{|I_P||j-m|^N} \lesssim \|f_3\|_p \frac{1}{|j-m|^N}.$$

In the event that $|j-m| \leq 2$, we have that

$$\left(\sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{|\langle f_3, \phi_P^1 \rangle|^2}{|I_P|} 1_{I_P} \right)^{1/2}$$

is a Littlewood–Paley-type mapping which is bounded from $L^s \to L^s$ for $s > 1$ (as shown, e.g., in [22]). Thus 5.2.6 is majorized by

$$\|f_3\|_p \min\left\{1, \frac{1}{|j-m|^N}\right\} \left\| \left(\sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \frac{|\langle g\chi_{I_{T,n}}, \tilde{\phi}_P^1 \rangle|^2}{|I_P|} 1_{I_P} \right)^{1/2} \right\|_{p'}$$

By employing a similar strategy, we see that the final factor in this last expression can be majorized by $|j-n|^{-N}$ as well.

Now, when $|j-n| \leq 2$, we need to estimate

$$\left\| \sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \chi_{I_{T,n}} \right\|_p,$$

where $j = n, n \pm 1, n \pm 2$. In such a case, $|j-m| \geq |m-n| - |j-n| \geq 5-2 = 3$. In such a case, one can drop the $\chi_{I_{T,n}}$, put the mods inside and make coarse estimates in a similar fashion as before to establish that

$$\left\| \sum_{P \in \mathcal{P}_{\mathbf{T}} : I_P \subseteq I_{T,j}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \chi_{I_{T,n}} \right\|_p \lesssim \|f_3\|_p \frac{1}{|j-m|^N}$$

In all, this produces, ultimately, that

$$\begin{aligned} \left\| \sum_{P \in \mathcal{P}_{\mathbf{T}}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \chi_{I_{T,n}} \right\|_p &\lesssim \sum_{j \in \mathbb{Z}} \|f_3\|_p \min\left\{1, \frac{1}{|j-n|^N}\right\} \min\left\{1, \frac{1}{|j-m|^N}\right\} \\ &= \|f_3\| \sum_{j \in \mathbb{Z}} \min\left\{1, \frac{1}{|j-n|^N}\right\} \min\left\{1, \frac{1}{|j-m|^N}\right\} \end{aligned}$$

Consider this last sum over $j \geq \frac{m+n}{2}$ and $j < \frac{m+n}{2}$ separately. We will focus on the first case as the other is very similar. Additionally, we assume that $m > n$ (recall $|m-n| \geq 5$) as the opposite case is no different. Moreover, there is no harm in assuming that $\frac{m+n}{2}$ is an integer (when $\frac{m+n-1}{2}$ is an integer it simply makes the notation more complicated). In such a situation the above sum easily transforms to

$$
\begin{aligned}
&\sum_{j=0}^{\infty} \min\left\{1, \frac{1}{|j - \frac{m-n}{2}|^N}\right\} \min\left\{1, \frac{1}{|j - \frac{n-m}{2}|^N}\right\} \\
&\leq \sum_{j=0}^{\infty} \min\left\{1, \frac{1}{|j - \frac{m-n}{2}|^N}\right\} \left(\frac{1}{|j| + |\frac{n-m}{2}|}\right)^N \\
&\leq \sum_{j=0}^{\infty} \min\left\{1, \frac{1}{|j - \frac{m-n}{2}|^N}\right\} \left(\frac{1}{|\frac{n-m}{2}|}\right)^N \\
&\lesssim \left(\frac{2}{|n-m|}\right)^N \\
&\lesssim \left(\frac{1}{|m-n|}\right)^N .
\end{aligned}
$$

Thus we have finally established our desired result when $|m-n| \geq 5$, namely

$$
\left\| \sum_{P \in \mathcal{P}_{\mathbf{T}}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \chi_{I_{T,n}} \right\|_p \lesssim \left(\frac{1}{|m-n|}\right)^N \|f_3\|_p.
$$

First term when $|m-n| < 5$. We now estimate the first term of (5.2.3) for $|m-n| < 5$. We again must split the sum over $\mathcal{P}_{\mathbf{T}}$ into a sum over those P with their time intervals in $I_{T,j}$ and then sum over j. If $|j-n| \gg 0$, then we may make cheap estimates as before, so we omit this case.

Thus we assume that $|j-n| \leq 5$, say. Each of these terms will be handled essentially identically. Now we drop the $\chi_{I_{T,n}}$, which is possible by positivity inside the norm. As stated just above, the P_1 tiles over any $I_{T,j}$ form a lacunary tree with respect to $I_{T,j} \times \omega_{T_3}$. Applying an identical argument to that in the latter part of estimating the third term of (5.2.3), we see that

$$\begin{aligned}\left|\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq I_{T,j}}\left\langle f_3,\phi_P^1\right\rangle\tilde{\phi}_P^1\right|& \\ &\leq\sup_{k\in\mathbb{Z}}\left|\left(\sum_{\mathcal{P}_{\mathbf{T}}:I_P\subseteq I_{T,j},|I_P|>2^k}\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right)\right| \\ &\lesssim M\left(\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq I_{T,j}}\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right),\end{aligned}$$

where M is the usual Hardy–Littlewood maximal operator. Since the Hardy–Littlewood maximal operator is bounded from $L^s\to L^s$ for $s>1$, we reduce to estimating

$$\left\|\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq I_{T,j}}\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right\|_p.$$

This can then be turned into a Littlewood–Paley square function as in the third term of (5.2.3) when $|n|\leq 2$. The boundedness of Littlewood–Paley square functions guarantees this is controlled by $\|f_3\|_p$.

We now put the previous results all together. Suppose that f_3 and f_4 are supported generically. Write $f_3=\sum_m f_3\chi_{I_{T.m}}$ and $f_4=\sum_n f_4\chi_{I_{T,n}}$. Applying the triangle inequality, we have, for the T which maximizes our $\mathrm{size}_3(\mathbf{T})$,

$$\begin{aligned}&\left|\frac{1}{|I_T|^{1/2}}\sum_{Q\in\mathbf{T}}a_Q\left\langle\sum_{P\in\mathcal{P}_{\mathbf{T}}}\left\langle f_3,\phi_P^1\right\rangle\tilde{\phi}_P^1\tilde{\phi}_Q^3,f_4\right\rangle\right| \\ &\lesssim\sum_{m,n\in\mathbb{Z}^2}\frac{1}{|I_2|^{1/2+1/2-1/r}}\frac{\|f_3\chi_{I_{T,m}}\|\|f_4\chi_{I_{T,n}}\|_q}{\max(1,|m-n|^N)\max(1,|n|^N)} \\ &=\sum_{m,n\in\mathbb{Z}^2}\frac{1}{|I_T|^{1-1/r}}\left\|f_3\frac{1}{|m|^{N/2}}\chi_{I_{T,m}}\right\|_p\left\|f_4\frac{1}{|n|^{N/2}}\chi_{I_{T,n}}\right\|_q\frac{\max(1,|m/n|^{N/2})}{\max(1,|m-n|^N)}\end{aligned}$$

If $|m|\leq|n|$, the last term is smaller than 1. If $m\in[kn,(k+1)n]$ for $|k|\geq 1$, then the last term is smaller than

$$\frac{(|k|+1)^{N/2}}{(\max\{1,|k|-1\})^N} \lesssim 1.$$

We may clearly pick a smooth, positive function $\tilde{\chi}_{I_T}$ which is adapted to I_T such that the function $\tilde{\chi}_{I_T}$ is larger than a constant times $|n|^{-N/2}$. Thus by disjointness of the intervals $I_{T,n}$ over n and $I_{T,m}$ over m,

$$\begin{aligned}
&\sum_{m,n\in\mathbb{Z}^2} \frac{1}{|I_T|^{1-1/r}} \left\| f_3 \frac{1}{|m|^{N/2}} \chi_{I_{T,m}} \right\|_p \left\| f_4 \frac{1}{|n|^{N/2}} \chi_{I_{T,n}} \right\|_q \\
&\lesssim \frac{1}{|I_T|^{1-1/r}} \left\| f_3 \sum_{m\in\mathbb{Z}} \frac{1}{|m|^{(N/2)-2}} \chi_{I_{T,m}} \right\|_p \left\| f_4 \sum_{n\in\mathbb{Z}} \frac{1}{|n|^{(N/2)-2}} \chi_{I_{T,n}} \right\|_q \\
&\lesssim \frac{\|f_3\tilde{\chi}_{I_T}\|_p \|f_4\tilde{\chi}_{I_T}\|_q}{|I_T|^{1-\frac{1}{r}}} \\
&= \frac{\|f_3\tilde{\chi}_{I_T}\|_p \|f_4\tilde{\chi}_{I_T}\|_q}{|I_T|^{1p+1/q}},
\end{aligned}$$

which is, finally, precisely what we wanted to show, provided we can justify the second line. But this simply follows by dualizing with some $g_m \in L^{p'}(I_{T,m})$ with p'-norm ≤ 1 and observing

$$\begin{aligned}
\sum_m \left\| f_3 \sum_{m\in\mathbb{Z}} \frac{1}{|m|^{N/2}} \chi_{I_{T,m}} \right\|_p &= \int \sum |f_3| \sum_{m\in\mathbb{Z}} \frac{1}{|m|^{(N/2)-2}} \chi_{I_{T,m}} \frac{1}{|m|^2} |g_m| dx \\
&\lesssim \int \sum |f_3| \sum_{m\in\mathbb{Z}} \frac{1}{|m|^{(N/2)-2}} \chi_{I_{T,m}} \sum_m \left(\frac{1}{1+|m|^2} g_m \right) dx \\
&\le \left\| \sum f_3 \sum_{m\in\mathbb{Z}} \frac{1}{|m|^{(N/2)-2}} \chi_{I_{T,m}} \right\|_p \left\| \sum_m \left(\frac{1}{|m|^2} g_m \right) \right\|_{p'} \\
&\lesssim \left\| \sum f_3 \sum_{m\in\mathbb{Z}} \frac{1}{|m|^{(N/2)-2}} \chi_{I_{T,m}} \right\|_p
\end{aligned}$$

□

This proposition is important for several reasons: first, it suggests that our main size estimate is potentially reasonable; second, it is a good "warm up" for understanding which tools are important; third, it will be the workhorse for the full estimate. We make the following observation:

Lemma 5.2.2. *Let $\overline{Q}$ be a fixed multitile in $\mathbf{T}$, with top $I_T \times \omega_T$. Let P be any tile in $\mathcal{P}_{\mathbf{T}}$. Then one of the following must occur:*

1. *$\omega_{P_2} \cap \omega_{\overline{Q}_3} \neq \emptyset$, in which case $\omega_{P_2} \supsetneq \omega_{\overline{Q}_3}$.*
2. *$\omega_{P_2} \cap \omega_{\overline{Q}_3} = \emptyset$, in which case ω_{P_2} lies closer to 0 than $\omega_{\overline{Q}_3}$ with $|\omega_{P_2}| < |\omega_{\overline{Q}_3}|$*

Proof. For (1), dyadicity guarantees that if the two intervals intersect one must contain the other. If $\omega_{P_2} \subseteq \omega_{\overline{Q}_3}$, then the fact that $P \in \mathcal{P}_{\mathbf{T}}$ means there is some $\omega_{Q'_3} \subsetneq \omega_{P_2}$ for some tile $Q' \in \mathbf{T}$. Hence there is a distinct tile Q' with $\omega_{Q'_3} \subsetneq \omega_{\overline{Q}_3}$ which violates the lacunarity property of $\mathbf{T}$.

For (2), first recall that all our trees either grow up or down, meaning that the ω_{Q_3} intervals all lie above or all lie below ω_{Q_T}. Observe that scale separation and the lacunarity property of $\mathbf{T}$ guarantees that if ω_{P_2} hits $\omega_{\overline{Q}_3}$ then ω_{P_2} is large enough that it contains ω_{Q_T} as well. Thus ω_{P_2} contains all the ω_{Q_3}'s lying between ω_{T_3} and $\omega_{\overline{Q}_3}$. Hence if $\omega_{P_2} \cap \omega_{\overline{Q}_3} = \emptyset$, then ω_{P_2} must have smaller frequency interval and lie closer to 0 than $\omega_{\overline{Q}_3}$.

$\square$

This lemma is quite useful as it allows us to establish the following corollary:

Corollary 5.2.3. *For each $Q \in \mathbf{T}$,*

$$\{P \in \mathbf{P} : \omega_{Q_3} \subsetneq \omega_{P_2}\} = \mathcal{P}_{\mathbf{T}} - \mathcal{P}_{Q,\text{lower}},$$

where $\mathcal{P}_{Q,\text{lower}}$ denotes the P_2 tiles from (2) in the previous lemma.

Before proceeding, we use linearity to split our operator into two pieces, one where $N_2(x) > N_1(x)$ and the other where $N_2(x) \leq N_1(x)$. We treat each case separately. The general strategy will be to perform some moves — which are similar to those we have made previously — to transform the sum over a maximal subtree into a discrete paraproduct.

5.2.1 Case I: $N_2(x) > N_1(x)$.

Proposition 5.2.4. *Suppose that $N_1(x)$ and $N_2(x)$ are integer-valued functions such that $N_2(x) > N_1(x)$ pointwise. Let f_3, f_4 be functions in $X(E_3)$ and $X(E_4)$ respectively, with $X(E)$ as defined at the beginning of this section. Suppose that $\tilde{\mathbf{T}}$ is a tree of Q-multitiles which is lacunary in the third position. Suppose further that $s \in \mathbb{Z}$ and*

$$2^s \leq \operatorname{size}_3(\tilde{\mathbf{T}}) \leq 2^{s+1}.$$

Then for any $\epsilon \in (0,1)$ and $p, q \in (1, \infty)$ with $1/p + 1/q = 1 - \epsilon$,

$$\operatorname{size}_3(\tilde{\mathbf{T}}) \lesssim \frac{\|f_3 \tilde{\chi}_{I_T}\|_p \|f_4 \tilde{\chi}_{I_T}\|_q}{|I_T|^{1/p+1/q}},$$

where $\mathbf{T}$ is the maximizing subtree for $\operatorname{size}_3(\tilde{\mathbf{T}})$ (which exists since $\tilde{\mathbf{T}}$ is a finite collection of multitiles).

Proof. Since all the nonzero terms in the sum have $|I_Q| > |I_P|$, we see that

$$|I_Q| > |I_P| > 2^{N_2(x)} > 2^{N_1(x)},$$

and so the factor of $1_{\{|I_Q| > 2^{N_1(x)}\}}$ is redundant and can be ignored in Case I. In particular, this means $\tilde{\phi}_Q^3 = \phi_Q^3$, and so this size estimate should be easier to manipulate than in the other case (where one does not immediately get this characteristic function to drop out). We proceed as in the previous proposition and let

$\mathbf{T}$ be the Q-tree which maximizes the quantity in the definition of $\text{size}_3(\tilde{\mathbf{T}})$. Then we need to estimate

$$\left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \sum_{P \in \mathbf{P}: \omega_{Q_3} \subsetneq \omega_{P_2}} \langle f_3, \phi_P^1 \rangle \tilde{\phi}_P^1 \phi_Q^3, f_4 \right\rangle \right|.$$

By invoking Corollary 5.2.3, we write this expression as

$$\left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \sum_{P \in \mathcal{P}_{\mathbf{T}}} \langle f_3, \phi_P^1 \rangle \tilde{\phi}_P^1 \phi_Q^3, f_4 \right\rangle \right|$$
$$+ \left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \sum_{P \in \mathcal{P}_{Q,\text{lower}}} \langle f_3, \phi_P^1 \rangle \tilde{\phi}_P^1 \phi_Q^3, f_4 \right\rangle \right|.$$

By Proposition 5.2.1, it suffices to deal only with the second term. Just as before, one can split f_3 and f_4 into the portions supported in $I_{T,m}$ and $I_{T,n}$, respectively, as well as turning the sum over $\mathcal{P}_Q$ into those P whose time intervals lie inside $I_{T,j}$; when m, n, j are not roughly equal, one makes coarse estimates as we did previously, putting the modulus inside the integrals, and so on, and get decay in m, n, j. The details for when $|m|, |n|, |j| \gg 0$ will be essentially identical as before, so we omit them. We will thus consider only those P whose time intervals lie within $5I_T$, say, and when f_m, f_n are supported in $3I_T$, in which case one cannot make such cheap estimates. So, we wish to establish the estimate

$$\left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \sum_{P \in \mathcal{P}_{Q,\text{lower}}: I_P \subseteq 5I_T} \langle f_3, \phi_P^1 \rangle \tilde{\phi}_P^1 \phi_Q^3, f_4 \right\rangle \right| \lesssim \frac{\|f_3\|_p \|f_4\|_q}{|I_T|^{1/p+1/q}},$$

where f_3, f_4 are supported in $3I_T$. We wish to change the order of summation; with that in mind, we define, for each $P \in \mathcal{P}_{\mathbf{T}}$ which contributes to the sum (i.e. for which there is some $Q \in \mathbf{T}$ such that $\omega_{Q_3} \subset \omega_{P_2}$),

$$\mathcal{Q}_{P,\text{upper}} := \{Q \in \mathbf{T} : |\omega_{Q_3}| > |\omega_{P_2}|\}.$$

Thus switching the order of the summation, we get

$$\begin{aligned}
&\left|\left\langle \sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T}\left(\sum_{Q\in\mathcal{Q}_{P,\text{upper}}} a_Q\phi_Q^3\right)\left\langle f_3,\phi_P^1\right\rangle\tilde{\phi}_P^1, f_4\right\rangle\right| \\
&\le \int\left|\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T}\left(\sum_{Q\in\mathcal{Q}_{P,\text{upper}}} a_Q\phi_Q^3\right)\left\langle f_3,\phi_P^1\right\rangle\tilde{\phi}_P^1\right||f_4(x)|dx \\
&\le \int\sup_k\left|\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T,|I_P|>2^k}\left(\sum_{Q\in\mathcal{Q}_{P,\text{upper}}} a_Q\phi_Q^3\right)\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right||f_4(x)|dx \\
&\le \int M\left(\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T}\left(\sum_{Q\in\mathcal{Q}_{P,\text{upper}}} a_Q\phi_Q^3\right)\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right)|f_4(x)|dx,
\end{aligned}$$

where M is the usual Hardy–Littlewood maximal operator. Applying Hölder's inequality and the boundedness of M, we reduce to simply estimating

$$\begin{aligned}
&\left\|\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T}\left(\sum_{Q\in\mathcal{Q}_{P,\text{upper}}} a_Q\phi_Q^3\right)\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right\|_{q'} \\
&=\left\|\sum_{Q\in\mathbf{T}} a_Q\left(\sum_{P\in\mathcal{P}_{Q,\text{lower}}:I_P\subseteq 5I_T}\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right)\phi_Q^3\right\|_{q'}.
\end{aligned}$$

Now, since the the P_1-tiles form a lacunary tree with respect to each $I_{T,j}\times\omega_{T_3}$, this inner sum can be written as a convolution of the full sum (i.e. over $\mathcal{P}_{\mathbf{T}}: I_P\subseteq 5I_T$) with a function whose Fourier transform is identically 1 on an interval which is roughly of length $|\omega_{Q_3}|$ and is 0 outside of a slightly larger interval. Write this function as $\varphi_{|Q|}$, which depends not on Q precisely but rater on the length of ω_{Q_3}. Thus we consider

$$\sum_{Q\in\mathbf{T}} a_Q\left(\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T}\left\langle f_3,\phi_P^1\right\rangle\phi_P^1\right)*\varphi_{|Q|}\phi_Q^3.$$

After distributing the convolution over the sum, we wish to discretize the convolution at scale $|I_Q|$; in the usual way, e.g. from [22], we write, for each $\alpha\in[0,1]$ and $i\in\mathbb{Z}$,

$$\varphi_{|Q|,i}^{\alpha}(x)=\overline{\varphi_{|Q|}(|I_Q|(i+\alpha)-x)}.$$

Observe that $\varphi^{\alpha}_{|Q|,i}$ is L^2-normalized. Then we have (as seen in [22])

$$\phi^1_P * \varphi_{|Q|} = \int_0^1 \sum_{i\in\mathbb{Z}} |I_Q|^{-1}\langle \phi^1_P, \varphi^{\alpha}_{|Q|,i}\rangle \varphi^{\alpha}_{|Q|,i} d\alpha.$$

This transforms our sum into

$$\int_0^1 \sum_i \sum_{Q\in\mathbf{T}} a_Q \left(\sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T} \left\langle f_3, \phi^1_P\right\rangle \left\langle \phi^1_P, \varphi^{\alpha}_{|Q|,i}\right\rangle \right) |I_Q|^{-1}\varphi^{\alpha}_{|Q|,i}\phi^3_Q d\alpha.$$

We will only estimate the inside of the integral with respect to α, which is fine provided the estimates do not depend on α (the dependence on α is harmless since it produces a complex exponential factor of controlled oscillation and the adaptedness can be made uniformly in α). The sum over i is certainly finite since the Fourier supports of ϕ^1_P and $\varphi^{\alpha}_{|Q|,i}$ are disjoint unless $i = 0, \pm 1$. By the triangle inequality, it suffices to consider each case separately; we treat each case identically, so we assume without loss of generality that $i = 0$. We now write $\Phi^{3,\alpha}_Q := |I_Q|^{-1/2}\varphi^{\alpha}_{|Q|,0}\phi^3_Q$. Observe that $\Phi^{3,\alpha}_Q$ is L^2 normalized since Parseval's identity and Young's inequality guarantee its L^2-norm is controlled by $|I_Q|^{-1/2}\|\widehat{\varphi^{\alpha}_{|Q|,0}}\|_1\|\widehat{\phi^3_Q}\|_2 \lesssim |I_Q|^{-1/2}|I_Q|^{1/2} = 1$. Its Fourier support is an interval which is roughly twice as long but shifted by $|\omega_{Q_3}|$. In particular, the $\Phi^{3,\alpha}_Q$ still represent a lacunary family of bumps. For ease of writing

$$F_3 := \sum_{P\in\mathcal{P}_{\mathbf{T}}:I_P\subseteq 5I_T} \langle f_3, \phi^1_P\rangle \phi^1_P.$$

Hence we need only estimate

$$\left\| \sum_{Q\in\mathbf{T}} \frac{1}{|I_Q|^{1/2}} a_Q \langle F_3, \varphi^{\alpha}_{|Q|,0}\rangle \Phi^{3,\alpha}_Q \right\|_{q'}.$$

Dualize with $g \in \text{Ł}^q$ and observe that this new object is like a discrete paraproduct on F_3, g, and the sequence a_Q which is BMO. With this in mind, we make standard moves:

$$
\begin{aligned}
&\left|\sum_{Q\in\mathbf{T}}\frac{1}{|I_Q|^{1/2}}a_Q\langle F_3,\varphi^{\alpha}_{|Q|,0}\rangle\langle g,\Phi^{3,\alpha}_Q\rangle\right| \\
&\leq \int\sum_{Q\in\mathbf{T}}\frac{|a_Q|}{|I_Q|^{1/2}}1_{I_Q}(x)\frac{|\langle F_3,\varphi^{\alpha}_{|Q|,0}\rangle|}{|I_Q|^{1/2}}1_{I_Q}(x)\frac{|\langle g,\Phi^{3,\alpha}_Q\rangle|}{|I_Q|^{1/2}}1_{I_Q}(x)dx \\
&\leq \int\sup_{I_Q:Q\in\mathbf{T}}\left(\frac{|\langle F_3,\varphi^{\alpha}_{|Q|,0}\rangle|}{|I_Q|^{1/2}}1_{I_Q}(x)\right)\sum_{Q\in\mathbf{T}}\frac{|a_Q|}{|I_Q|^{1/2}}1_{I_Q}(x)\frac{|\langle g,\Phi^{3,\alpha}_Q\rangle|}{|I_Q|^{1/2}}1_{I_Q}(x)dx.
\end{aligned}
$$

First, observe that since $|I_Q|^{-1/2}\varphi^{\alpha}_{|Q|,0}$ is a nice L^1-normalized non-lacunary sequence of bump function, the first term here is controlled by $M(F_3)$. Applying the Cauchy–Schwarz inequality (with respect to the sum on the first two terms) and then Hölder's inequality (into L^r, L^p, and L^q appropriately), one must estimate

$$
\|M(F_3)\|_p\left\|\left(\sum_{Q\in\mathbf{T}}\frac{|a_Q|^2}{|I_Q|}1_{I_Q}(x)\right)^{1/2}\right\|_r\left\|\left(\sum_{Q\in\mathbf{T}}\frac{|\langle g,\Phi^{3,\alpha}_Q\rangle|^2}{|I_Q|}1_{I_Q}(x)\right)^{1/2}\right\|_q.
$$

One then estimates these precisely as we have in the proof of Proposition 5.2.3 to get the desired result (which are of course independent of α).

□

5.2.2 Case II. $N_2(x)\leq N_1(x)$.

Proposition 5.2.5. *Suppose that $N_1(x)$ and $N_2(x)$ are integer-valued functions such that $N_2(x)\leq N_1(x)$, pointwise. Let f_3, f_4 be functions in $X(E_3)$ and $X(E_4)$ respectively, with $X(E)$ as defined at the beginning of this section. Suppose that $\tilde{\mathbf{T}}$ is a tree of Q-multitiles which is lacunary in the third position. Suppose further that $s\in\mathbb{Z}$ and*

$$2^s\leq \mathrm{size}_3(\tilde{\mathbf{T}})\leq 2^{s+1}.$$

Then for any $\epsilon \in (0,1)$ *and* $p, q \in (1, \infty)$ *with* $1/p + 1/q = 1 - \epsilon$*,*

$$\text{size}_3(\tilde{\mathbf{T}}) \lesssim \frac{\|f_3 \tilde{\chi}_{I_T}\|_p \|f_4 \tilde{\chi}_{I_T}\|_q}{|I_T|^{1/p+1/q}},$$

where $\mathbf{T}$ *is the maximizing subtree for* $\text{size}_3(\tilde{\mathbf{T}})$ *(which exists since* $\tilde{\mathbf{T}}$ *is a finite collection of multitiles).*

Proof. The main argument here is nearly identical to that of Proposition 5.2.4, with the caveat that one does not get the characteristic function to drop out at the start. However, when one moves to the sum over $\mathcal{P}_{Q,\text{lower}}$, one observes that all the tiles involved in this sum have

$$|I_P| > |I_Q| > 2^{N_1(x)} \geq 2^{N_2(x)}.$$

Thus in this sum, the characteristic function, $1_{|I_P| > 2^{N_2(x)}}$ drops out, and we find $\tilde{\phi}_P^1 = \phi_P^1$. So we need to estimate

$$\left| \frac{1}{|I_T|^{1/2}} \sum_{Q \in \mathbf{T}} a_Q \left\langle \sum_{P \in \mathcal{P}_{Q,\text{lower}}} \left\langle f_3, \phi_P^1 \right\rangle \phi_P^1 \tilde{\phi}_Q^3, f_4 \right\rangle \right|$$

However, one can again majorize the by as

$$\int \left| \sum_{Q \in \mathbf{T}} a_Q \left(\sum_{P \in \mathcal{P}_{Q,\text{lower}}} \langle f_3, \phi_P^1 \rangle \phi_P^1 \right) \tilde{\phi}_Q^3 \right| |f_4(x)| dx.$$

But now one can apply Hölder's inequality and majorize this sum with the characteristic function by the Hardy–Littlewood of the sum without the characteristic function. The proof from here on is then precisely identical to that of Proposition 5.2.4.

□

CHAPTER 6

HEURISTIC ARGUMENT FOR ENERGIES FOR $M = 0$, $K_2 \leq K_1$

In this chapter, we provide a terse heuristic argument for producing the energy estimate. We will omit most of the details but provide the reader with a high-level overview of our aims; a rigorous argument will need to be formulated carefully later on. In particular, we ignore the presence of the $b_{P,Q}$ factor which appears; as before, one would have to write it as a Taylor series whose terms are nicely behaved on trees. We will reduce the relevant result to certain estimates over trees, so this should not be an issue. In any case, the "main term" from the previous chapter, i.e. the $m' = 0$ term from the second Taylor series, would correspond to the one we discuss here.

6.1 Main Goal

Our main goal will be to produce an analogue for the workhorse theorem of Demeter, Tao, and Thiele's paper, [8]:

$$\sum_{\vec{P} \in \cup_{T \in \mathcal{F}} T} |\langle f_4, \phi_P^3 1_{|I_P| > 2^{N(x)}} \rangle|^2 \lesssim \|N_{\mathcal{F}}\|_\infty^{1/\mu} \int |f_4|^2 \chi_{I_0}^{10},$$

where N is an arbitrary integer-valued function and $\mathcal{F}$ is a collection of disjoint 3-lacunary trees which additionally satisfy the conditions on size of Lemma 2.2.4. In particular, we would like to come up with an estimate of the following form for a forest $\mathcal{F}$:

$$\sum_{Q \in \cup_{T \in \mathcal{F}} T} \left| \left\langle \tilde{\phi}_Q^3, \sum_{\omega_{Q_3} \subsetneq \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right\rangle \right|^2 \lesssim \|N_{\mathcal{F}}\|_\infty^{1/\mu} \|f_3\|_p \|f_4\|_q$$

where $1/p + 1/q = 1/2$ (or at least very close to $1/2$). Additionally, the forest $\mathcal{F}$ must also satisfy the size constraints of Lemma 2.2.4.

6.2 Model and First Heuristic

Our original model is the following:

$$\sum_{Q} \frac{1}{|I_Q|^{1/2}} \left\langle f_1, \phi_Q^1 \right\rangle \left\langle f_2, \phi_Q^2 \right\rangle \left\langle \tilde{\phi}_Q^3, \sum_{\omega_{Q_3} \subsetneq \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right\rangle .$$

Since the 1-energy and 2-energy of this expression are standard, we are left to estimate the 3-energy of this expression, i.e. to estimate

$$\sum_{Q \in \mathcal{D}} \left| \left\langle \tilde{\phi}_Q^3, \sum_{\omega_{Q_3} \subsetneq \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right\rangle \right|^2 ,$$

where $\mathcal{D}$ is some family of of multitiles whose 3rd position subtiles are disjoint. Linearize this expression by dualizing with a sequence a_Q in ℓ^2, and recall that restricting a_Q to a family of tiles in $\mathcal{D}$ which make up a lacunary 3-tree whose size is between 2^s and 2^{s+1} results in a sequence with BMO-norm controlled by $|I_T|^{-1/2}$, where I_T Is the time interval for the top of the maximizing (with respect to size) subtree. Thus we study

$$\sum_{Q \in \mathcal{D}} a_Q \left\langle \tilde{\phi}_Q^3, \sum_{\omega_{Q_3} \subsetneq \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right\rangle .$$

6.2.1 Heuristic Argument

Write the above expression as

$$\sum_{Q \in \mathcal{D}} a_Q \sum_{\omega_{Q_3} \subseteq \omega_{P_2}} \left\langle \tilde{\phi}_Q^3, \tilde{\phi}_P^1 f_4 \right\rangle \left\langle f_3, \phi_P^1 \right\rangle .$$

Suppose for the moment that the characteristic function requiring $|I_Q| > 2^{N_1(x)}$ dropped out, i.e. we had no tilde over ϕ_Q^3. For a fixed Q and P, discretize ϕ_Q^3

according to the scale $|I_P|$ and frequency interval ω_{P_2}. More specifically, let φ_P^2 be an L^1-normalized function whose Fourier transform is identically 1 on ω_{P_2}, 0 outside $1.2\omega_{P_2}$ and which satisfies the usual rapid decay estimates away from I_P in units of $|I_P|$. Heuristically speaking, one should think of φ_P^2 as being adapted to the tile $I_P \times \omega_{P_2}$. Now, we write, as in [22], for each $\alpha \in [0,1]$ and $j \in \mathbb{Z}$,

$$\varphi_{P,j}^{2,\alpha}(x) = |I_P|^{1/2}\overline{\varphi_P^2(|I_P|(j+\alpha) - x)}.$$

Thus $\varphi_{P,j}^{2,\alpha}$ has the same Fourier support while "morally" supported on an interval of length $|I_P|$ which is translated by $j+\alpha$ units of length $|I_P|$ — also note that $\varphi_{P,j}^{2,\alpha}$ is L^2-normalized. We have, for all P, Q involved in our sum, $\widehat{\phi_Q^3} = \widehat{\phi_Q^3}\widehat{\varphi_P^2}$, and hence $\phi_Q^3 = \phi_Q^3 * \varphi_P^2$. Then we can write, as in [22],

$$\left\langle \phi_Q^3, \tilde{\phi}_P^1 f_4 \right\rangle = \int_0^1 \sum_{j\in\mathbb{Z}} \left\langle \langle \phi_Q^3, \varphi_{P,j}^{2,\alpha}\rangle \varphi_{P,j,2}^{2,\alpha}, \tilde{\phi}_P^1 f_4 \right\rangle d\alpha,$$

where $\varphi_{P,j,2}^{2,\alpha}$ is morally the same function as $\varphi_{P,j}^{2,\alpha}$ except that it has a slightly larger Fourier support (on $1.4\omega_{P_2}$, say). The energy can now be written as

$$\int_0^1 \sum_{j\in\mathbb{Z}} \sum_{Q\in\mathcal{D}} a_Q \sum_{\omega_{Q_3} \subsetneq \omega_{P_2}} \langle \phi_Q^3, \varphi_{P,j}^{2,\alpha}\rangle \left\langle \varphi_{P,j,2}^{2,\alpha}, \tilde{\phi}_P^1 f_4 \right\rangle \langle f_3, \phi_P^1\rangle \, d\alpha.$$

Owing to the extreme decay of ϕ_P^1 away from I_P, we expect only the $j=0$ term to significantly contribute to the sum (or, more accurately, just a few terms besides $j=0$), while α is also mostly inconsequential for the same reasons given in earlier sections.

We now write $\Phi_P^2 := \varphi_{P,0}^{2,\alpha}$ and $\Phi_P^3 := |I_P|^{1/2}\varphi_{P,0}^{2,0}\phi_P^1$. Observe that Φ_P^2 and Φ_P^3 are both L^2-normalized; moreover, $\widehat{\Phi_P^3}$ is morally supported on $\omega_{P_1} + \omega_{P_2}$. So, the energy can be written as

$$\sum_{Q\in\mathcal{D}} a_Q \sum_{\omega_{Q_3} \subset \omega_{P_2}} \frac{1}{|I_P|^{1/2}} \langle f_3, \phi_P^1\rangle \langle \phi_Q^3, \Phi_P^2\rangle \left\langle f_4, \widetilde{\Phi}_P^3 \right\rangle,$$

which can be rearranged to

$$\sum_P \frac{1}{|I_P|^{1/2}} \left\langle f_3, \phi_P^1 \right\rangle \left\langle \sum_{\omega_{Q_3} \subset \omega_{P_2}} a_Q \phi_Q^3, \Phi_P^2 \right\rangle \left\langle f_4, \widetilde{\Phi}_P^3 \right\rangle .$$

This is then itself a trilinear form. The first term has standard size and energy estimates. The last term has size and energy estimates available in the the spirit of Demeter, Tao, Thiele, i.e. [8]. The middle term can then be estimated using Biest-type estimates — in particular, we can produce energy-type estimates, but *no size estimate is available.*

We remark that, for the non-simplified model, we will again split into two separate cases, where $N_2(x) \geq N_1(x)$ and $N_1(x) > N_2(x)$, respectively. In the former situation, the tilde over ϕ_Q^3 will drop out as it did with the sizes, and we may directly apply the argument described in this heuristic. In the latter situation, we will carefully carve the full collection of tiles into subsets with an appropriate tree structure; we will use a combination of the Rademacher–Menshov theorem and a modified lemma of Bourgain (from [4]) in a similar way as presented in [8] to reduce the problem to making estimates on individual trees, modulo a certain logarithmic loss. We have already produced estimates in trees, and the logarithmic loss, it will turn out, is tolerable.

6.3 Application to model when $N_2(x) \geq N_1(x)$

Observe that in this case, for every pair Q, P which participates in the sum,

$$|I_Q| > |I_P| > 2^{N_2(x)} \geq 2^{N_1(x)},$$

and so the constraint on $\tilde{\phi}_Q^3$ that $|I_Q| > 2^{N_1(x)}$ is redundant and disappears. Thus we may directly apply the heuristic from the previous section.

6.4 Application to model when $N_1(x) > N_2(x)$

In this case, both constraints appear. Dualize as before and write the expression as

$$\int \sum_{Q\in\mathcal{D}} a_Q \tilde{\phi}_Q^3 \sum_{\omega_{Q_3}\subset\omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 dx.$$

Owing to a pointwise estimate on the above integrand, this can be majorized by replacing the tilde over ϕ_Q^3 by a supremum:

$$\int \sup_k \left| \sum_{Q\in\mathcal{D},|I_Q|>2^k} a_Q \phi_Q^3 \sum_{\omega_{Q_3}\subset\omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right| dx.$$

As a consequence of the heuristic discussion in the previous section, the term inside the supremum can be estimated using known techniques.

6.5 Splitting the sum

Following the techniques of [8], we split our family of trees $\mathcal{Q}$ into two parts. We define the following set,

$$\mathcal{I} := \{I : I_Q \text{ is the time interval of the top of a tree in } \mathcal{Q}\}.$$

We now construct a sequence $\mathcal{I}_j$ of sets in $\mathcal{I}$, selecting by order of inclusion (i.e. define the sets recursively, picking the maximal intervals at each step). Observe that there are exactly $\|N_{\mathcal{Q}}\|_\infty$ distinct $\mathcal{I}_j$. Now define

$$\mathcal{Q}_{I,j} := \{Q \in \mathcal{Q} : I_Q = I \text{ where } I \in \mathcal{I}_j\}$$

and

$$\mathcal{Q}_{<I,j} := \{Q \in \mathcal{Q} : I_Q \subset I \text{ with } I \in \mathcal{I}_j \text{ but } I_Q \nsubseteq \mathcal{I}_l \text{ for any } l > j\}.$$

One may now clearly split the expression from the previous section into the sum of

$$\int \sup_k \left| \sum_j \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{I,j}, |I_Q| > 2^k} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right| dx. \tag{6.5.1}$$

and

$$\int \sup_k \left| \sum_j \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{<I,j}, |I_Q| > 2^k} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right| dx. \tag{6.5.2}$$

We may focus on these terms separately. The first lends itself to a variant of the Rademacher-Menshov theorem and the second relates to a variant of a lemma of Bourgain.

6.5.1 Rademacher–Menshov and Bourgain lemmas

Theorem 6.5.1 (Rademacher-Menshov)**.** *Let f_i for $i \in \{1, ..., L\}$ be a sequence of functions in $L^2(\mathbb{R})$ which are almost orthogonal. Specifically, for any sequence $\epsilon_1, ..., \epsilon_L \in \{-1, 1\}$, we have*

$$\left\| \sum_{i=1}^{L} \epsilon_i f_i \right\|_2 \leq B.$$

Then

$$\left\| \sup_{L' \leq L} \left| \sum_{i=1}^{L'} f_i \right| \right\|_2 \lesssim B \log(2 + L).$$

Proof. See [8, Theorem 10.6] □

The following is [4, Lemma 4.11] translated into our notation:

Theorem 6.5.2 (Bourgain's Lemma)**.** *Let Π_k be a Fourier projection onto the union of J intervals centered at $\xi_1, ..., \xi_J$, each of length $\sim 2^{-k}$. Then*

$$\left\| \sup_k |\Pi_k f| \right\|_2 \lesssim \log(2 + J)^2 \|f\|_2.$$

Proof. Bourgain's proof in [4, Lemma 4.11] is quite elegant and somewhat reminiscent, although significantly more abstract, of the proof of the Rademacher–Menshov Theorem from [8]. □

6.5.2 The first term and Rademacher-Menshov

In this section we discuss the first term, (6.5.1), i.e.

$$\int \sup_k \left| \sum_j \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{I,j}, |I_Q| > 2^k} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right| dx.$$

Since j essentially controls the scale of the term involved, it is not hard to see that, modulo an error term (as in the estimation of (60) from [8, Section 11]), the constraint on k can be shifted to a constraint on j so that we can majorize this expression by

$$\int \sup_{j_0} \left| \sum_{j \le j_0} \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{I,j}} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right| dx.$$

Apply Cauchy-Schwarz to reduce to estimating

$$\left\| \sup_{j_0} \sum_{j \le j_0} \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{I,j}} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \right\|_2 .$$

The Rademacher-Menshov theorem given above now applies directly, and this may be bounded by

$$\log\left(2 + \|N_{\mathcal{Q}}\|_\infty\right) \left\| \sum_j \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{I,j}} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 \right\|_2 .$$

Now one may proceed as in the heuristic provided above since this last quantity may be estimated using Biest-type arguments.

6.5.3 The second term and Bourgain's lemma

In this section, we discuss the second term, (6.5.2), i.e.

$$\int \sup_k \left| \sum_j \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{<I,j}, |I_Q| > 2^k} a_Q \phi_Q^3 \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4 \right| dx.$$

For ease of notation, we write

$$m_Q(f_3, f_4) := a_Q \sum_{\omega_{Q_3} \subset \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 f_4.$$

We consider the expression inside the supremum for some fixed k. Fix x and assume x is not the endpoint of a dyadic interval. We see that

$$\begin{aligned}
&\left| \sum_j \sum_{I \in \mathcal{I}_j} \sum_{Q \in \mathcal{Q}_{<I,j}, |I_Q| > 2^k} m_Q(f_3, f_4) \phi_Q^3 \right| \\
&\leq \left| \sum_j \sum_{I \in \mathcal{I}_j; x \in I} \sum_{Q \in \mathcal{Q}_{<I,j}, |I_Q| > 2^k} m_Q(f_3, f_4) \phi_Q^3 \right| \\
&+ \left| \sum_j \sum_{I \in \mathcal{I}_j; x \notin I} \sum_{Q \in \mathcal{Q}_{<I,j}, |I_Q| > 2^k} m_Q(f_3, f_4) \phi_Q^3 \right| \\
&:= I + II.
\end{aligned}$$

One should think of the first term as the main term and the second term as an error term; we will not discuss the second term here.

The first term, I

Here we estimate

$$\left| \sum_j \sum_{I \in \mathcal{I}_j; x \in I} \sum_{Q \in \mathcal{Q}_{<I,j}, |I_Q| > 2^k} m_Q(f_3, f_4) \phi_Q^3 \right|.$$

Observe that there is a largest $j_0 = j_0(x)$ such that such that there is some interval in $\mathcal{I}_{j_0}$ containing x which has length greater than 2^k. Only one such interval exists, denote it I_0. Hence

$$\left|\sum_j \sum_{I\in\mathcal{I}_j;x\in I} \sum_{Q\in\mathcal{Q}_{<I,j},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right|$$
$$\le \left|\sum_{j<j_0(x)} \sum_{I\in\mathcal{I}_j;x\in I} \sum_{Q\in\mathcal{Q}_{<I,j},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right| + \left|\sum_{Q\in\mathcal{Q}_{<I_0,j_0},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right|$$
$$\le \left|\sum_{j<j_0(x)} \sum_{I\in\mathcal{I}_j} \sum_{Q\in\mathcal{Q}_{<I,j},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right| + \left|\sum_{Q\in\mathcal{Q}_{<I_0,j_0},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right|$$

In the first term, the condition that $j < j_0$ forces all the I in $\mathcal{I}_j$ to have length greater 2^k, and so this condition is redundant. One can drop the condition that j_0 depends on x by replacing the first term with the supremum over all j_0 of the first term. This can then be handled using the Rademacher-Menshov theorem exactly as in the previous section. Thus we focus on the other term, which is clearly majorized by

$$\sup_{I\in\mathcal{I}} \sup_k \left|\sum_{Q\in\mathcal{Q}_{<I},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right|.$$

The obvious inequalities

$$\|\sup_i |f_i|\|_2 \le \left\|\left(\sum_i |f_i|^2\right)^{1/2}\right\|_2 = \left(\sum_i \|f_i\|_2^2\right)^{1/2}$$

now leave us to prove

$$\left(\sum_{I\in\mathcal{I}} \left\|\sup_k \left|\sum_{Q\in\mathcal{Q}_{<I},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right|\right\|_2^2\right)^{1/2} \lesssim \log(2+\|N_{\mathcal{Q}}\|_\infty)^2\|f_3\|_p\|f_4\|_q\|a_Q\|_{\ell^2}$$

for appropriate p, q, which of course would follow by

$$\left\|\sup_k \left|\sum_{Q\in\mathcal{Q}_{<I},|I_Q|>2^k} m_Q(f_3,f_4)\phi_Q^3\right|\right\|_2 \lesssim \log(2+\|N_{\mathcal{Q}}\|_\infty)^2\|f_3\|_p\|f_4\|_q \left(\sum_{Q_{<I}} |a_Q|^2\right)^{1/2}$$

for each I.

To that end, let $\mathbf{T}_1, ..., \mathbf{T}_J$ denote the trees in $\mathcal{Q}$ which intersect $P_{<I}$. Necessarily, $J \leq \|N_{\mathcal{Q}}\|_\infty$. Thus we write

$$\sum_{Q \in \mathcal{Q}_{<I}, |I_Q| > 2^k} m_Q(f_3, f_4)\phi_Q^3 = \sum_{j=1}^{J} \sum_{Q \in \mathcal{Q}_{<I} \cap \mathbf{T}_j, |I_Q| > 2^k} m_Q(f_3, f_4)\phi_Q^3.$$

It is not hard to see that, from the disjointness of the trees, one can rewrite this as a Fourier projection onto the union of J intervals centered at $\xi_1, ..., \xi_J$, the centers of the frequency intervals of the tops of the trees. One would like to apply Bourgain's lemma directly, but this is a bit delicate. For one, the P tiles may have very large supports, and so one cannot write this as a sum of Fourier projections. However, since we are working on trees, we have some tricks available. On each tree, we again write

$$\sum_{\omega_{Q_3} \subsetneq \omega_{P_2}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 = \sum_{P \in \mathcal{P}_{T_j}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1 - \sum_{\omega_{P_2} \text{ below}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1.$$

The second term concentrates the P-tiles onto frequencies near the top of the Q-tree. Moreover, for every P, Q which contribute to the sum below,

$$|I_P| > |I_Q| > 2^{N_1(x)} > 2^{N_2(x)},$$

and so we know that the constraint on P can be dropped in this case. Thus we may exactly write this term as a Fourier projection and apply Bourgain's lemma exactly as stated, then use the Biest-type estimates once we have gotten rid of the supremum over k.

We now focus on the first term. Write

$$\tilde{F}_j(x) = \sum_{P \in \mathcal{P}_{T_j}} \left\langle f_3, \phi_P^1 \right\rangle \tilde{\phi}_P^1(x),$$

and $F_j(x)$ for the same term but without the tilde over ϕ_P^1. Observe that we have the benefit that $\tilde{F}_j$ no longer depends on Q. One can carefully read Bourgain's proof of Lemma 4.13 from [4] (which is really all we need) and find that his theorem is completely generic provided one can produce estimates of

$$\left\| \tilde{F}_j(x) f_4(x) \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} a_Q \phi_Q^3 \right) * \phi_t \right\|_{v_s} \right\|_2^2 ,$$

for s arbitrarily close to 2. Since, morally, all the P tiles in F_j should be supported on the top of the Q-tree, apply Hölder to see, roughly speaking, that for $p^{-1}+q^{-1} = 1/2$,

$$\begin{aligned} &\left\| \tilde{F}_j(x) f_4(x) \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} a_Q \phi_Q^3 \right) * \phi_t \right\|_{v_s} \right\|_2 \\ &\leq \left\| \tilde{F}_j(x) f_4(x) \right\|_{L^p(10 I_{T_j})} \left\| \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} a_Q \phi_Q^3 \right) * \phi_t \right\|_{v_s} \right\|_q . \end{aligned}$$

Now, by an obvious pointwise estimate, one has

$$\begin{aligned} &\left\| \tilde{F}_j(x) f_4 \right\|_{L^p(10 I_{T_j})} \leq \left\| \tilde{F}_j(x) \right\|_{L^{p_1}(10 I_{T_j})} \| f_4 \|_{L^{p_2}(10 I_{T_j})} \\ &\lesssim \| M(F_j) \|_{L^{p_1}(10 I_{T_j})} \| f_4 \|_{L^{p_2}(10 I_{T_j})} \lesssim \| F_j \|_{L^{p_1}(10 I_{T_j})} \| f_4 \|_{L^{p_2}(10 I_{T_j})}, \end{aligned}$$

for $p_1^{-1} + p_2^{-1} = p^{-1}$. For the other term, observe that, for large $q > 2$, one should have

$$\left\| \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} a_Q \phi_Q^3 \right) * \phi_t \right\|_{v_s} \right\|_q \lesssim \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} a_Q \phi_Q^3 \right) \right\|_q ,$$

which is an analogue of the main results in [29] (in that paper, the function ϕ_t would have Fourier transform equal to a characteristic function of an interval, i.e.

which is not smooth). Dualize with a function g in $L^{q'}$ with $\|g\|_{q'} = 1$:

$$\begin{aligned}
&\left\| \left(\sum_{Q \in Q_{<I} \cap T_j} a_Q \phi_Q^3 \right) \right\|_q = \sum_{Q \in Q_{<I} \cap T_j} a_Q \left\langle g, \phi_Q^3 \right\rangle \\
&= \int \sum_{Q \in Q_{<I} \cap T_j} \frac{a_Q}{|I_Q|^{1/2}} 1_{I_Q}(x) \left\langle g, \phi_Q^3 \right\rangle \frac{1}{|I_Q|^{1/2}} 1_{I_Q}(x) dx \\
&\le \int \left(\sum_{Q \in Q_{<I} \cap T_j} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \left(\sum_{Q \in Q_{<I} \cap T_j} \left\langle g, \phi_Q^3 \right\rangle \frac{1}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} dx \\
&\lesssim \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_q \left\| \left(\sum_{Q \in Q_{<I} \cap T_j} \left\langle g, \phi_Q^3 \right\rangle \frac{1}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_{q'}
\end{aligned}$$

The second term is a square function of g, and so its q'-norm is majorized by 1. By using the same discretized BMO norm described in the previous chapter,

$$\left\| \left(\sum_{Q \in Q_{<I} \cap T_j} \frac{|a_Q|^2}{|I_Q|} 1_{I_Q}(x) \right)^{1/2} \right\|_q \lesssim |I_T|^{1/p} \|a_Q\|_{BMO(I_{T_j})}.$$

By virtue of the properties of the sequence a_Q from our selection algorithm, we of course have that the length of the top controls the BMO-norm of the sequence restricted to a tree:

$$\|a_Q\|_{BMO(I_{T_j})} \lesssim |I_T|^{-1/2} \|a_Q\|_{\ell^2(Q \in Q_{<I} \cap T_j)},$$

which means that this whole term is majorized by

$$\begin{aligned}
&\|a_Q\|_{\ell^2(Q \in Q_{<I} \cap T_j)} |I_T|^{-1/2+1/q} \|F_j\|_{L^{p_1}(10 I_{T_j})} \|f_4\|_{L^{p_2}(10 I_{T_j})} \\
&= \|a_Q\|_{\ell^2(Q \in Q_{<I} \cap T_j)} \frac{\|F_j\|_{L^{p_1}(10 I_{T_j})} \|f_4\|_{L^{p_2}(10 I_{T_j})}}{|I_T|^{1/p}},
\end{aligned}$$

with $p_1^{-1} + p_2^{-1} = p^{-1}$, where p can be taken as close to 2 as we'd like. It can easily be seen that the P-tiles in the sum in F_j inherit the tree structure (when restricted to a tree top I_T), so $\|F_j\|_{p_1}$ can be controlled by $\|f_3\|_{p_1}$ and so this last part is a a product of averages of f_3 and f_4. This completes the argument for term **I**.

Now, one should be able to proceed essentially as in [8, Chapter 9], with this result as a replacement for Theorem 9.2.

CHAPTER 7

FURTHER WORK

The beauty of the time-frequency approach to maximal multilinear operators (inspired by Lacey, [17], and Demeter, Tao, and Thiele, [8]) is that one simultaneously establishes estimates on (maximal) singular integrals. The benefit of such a theory is manifest: for example, the "simplest" singular integral operator, the Hilbert transform, can only be defined explicitly on a dense class of functions and then extended to all of L^2 by a norm-limit. The maximal singular integral estimate says that, in fact, this limiting argument also holds pointwise — thus one can actually visualize the graph of the Hilbert transform of a generic L^2 function. That said, the power of the techniques comes at a price: certain maximal operators are bounded while the related singular integral is known to diverge (and hence the time-frequency method cannot be used to say anything about the maximal operator); even when the singular integrals are bounded, the estimates provable using this method are fundamentally limited while the related maximal operators seem to be bounded in a broader range. We state some open questions which arise naturally from the ideas surrounding the present work.

7.1 Further work

There are a number of natural questions which arise from the discussion above. We conclude this thesis by stating several of them.

Question 1. *Do the Taylor series terms for $m \geq 1$, $m' \geq 1$ discussed in this text satisfy the same estimates as the remainder and $m = 0$, $m' = 0$ terms? This would* *follow by thesis work currently being carried out by Joeun Jung, another of Camil Muscalu's students.*

Question 2. *One immediately natural is to complete the general problem which inspired my thesis: that is, to determine whether or not the ergodic average in (0.2.1) converges pointwise over* $L^\infty \times L^\infty \times L^\infty$. *This would follow, e.g., by establishing bi-parameter variational-type estimates in the spirit of Oberlin, Seeger, Tao, Thiele, and Wright, [29], and Do, Muscalu, and Thiele, [9], which use the tools of time-frequency analysis. In many ways the whole previous discussion is a "warm up" for these more intricate estimates.*

Question 3. *The bilinear maximal operator described above, (0.3.1),* ***for a fixed*** r *(i.e. without the* $\sup$*), is bounded from* $L^p \times L^q \to L^s$ *for* $s^{-1} = p^{-1} + q^{-1}$ *for* s *all the way to* $1/2$. *However, the singular integral techniques employed are known to break down at* $s = 2/3$. *Can we get estimates for this bilinear maximal operator for* s *closer to* $1/2$? *In using Fourier analysis, we have made the estimates much simpler to achieve, but we dropped the positivity condition on our functions, so this may indeed be possible.*

Question 4. *In (0.2.2) above, if one replaced* $f(x+s+t)$ *with* $f(x-s-t)$ *and forced* $h_1 = h_2$, *the related one-parameter singular integral operator will not be bounded for* $f_i \in L^2$ *even though the singular integral operator related to the original operator is bounded for* $f_i \in L^2$. *However, the corresponding (one-parameter) ergodic averages are known to converge pointwise over* $L^\infty \times L^\infty \times L^\infty$ *by Assani, [1], indicating that this modified maximal operator probably satisfies non-trivial estimates. Can we prove that some of these estimates are achievable? If not, can it tell us something about singular integrals?*

Question 5. *In the discussion of time-frequency analysis and AKNS systems provided in the introduction, we mentioned that Christ and Kiselev were interested* *in proving the boundedness of orbits for Schrödinger operators with L^2 potentials. This is presumably still an open question. Additionally, we may pose the related question for general AKNS systems: are orbits bounded when the V_{ij} are in L^2? This is open, even when $n = 2$.*

BIBLIOGRAPHY

[1] I. Assani. Pointwise convergence of ergodic averages along cubes. *Journal D'Analyse Mathématique*, 110:241–269, 2010.

[2] T. Austin. On the norm convergence of non-conventional ergodic averages. *Ergodic Theory and Dynamical Systems*, 30(2):321–338, 2010.

[3] C. Bennett and R. Sharpley. *Interpolation of operators*, volume 129 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1988.

[4] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.*, (69):5–45, 1989.

[5] J. Bourgain. Double recurrence and almost sure convergence. *J. Reine Angew. Math.*, 404:140–161, 1990.

[6] M. Christ and A. Kiselev. WKB and Spectral Analysis of One-Dimensional Schrödinger Operators with Slowly Varying Potentials. *Comm. Math. Phys.*, 218(2):245–262, 2001.

[7] M. Christ and A. Kiselev. WKB Asymptotic Behavior of Almost All Generalized Eigenfunctions for One-Dimensional Schrödinger Operators with Slowly Decaying Potentials. *J. Funct. Anal.*, 179(2):426–447, 2001.

[8] C. Demeter, T. Tao, and C. Thiele. Maximal multilinear operators. *Trans. Amer. Math. Soc.*, 360(9):4989–5042, 2008.

[9] Y. Do, C. Muscalu, and C. Thiele. Variational estimates for paraproducts. *Rev. Mat. Iberoam.*, 28(3):857–878, 2012.

[10] C. Fefferman. Pointwise convergence of Fourier series. *Ann. of Math. (2)*, 98:551–571, 1973.

[11] C. Fefferman. The Uncertainty Principle. *Bull. Amer. Math. Soc. (N.S.)*, 9(2):129–206, 1983.

[12] H. Furstenberg and B. Weiss. A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$. In V. Bergleson, A. March, and J. Rosenblatt, editors, *Convergence in Ergodic Theory and Probability*, volume 5 of *Ohio State Univ. Math. Res. Inst. Publ.*, pages 193–227. de Gruyter, Berlin, 1996.

[13] L. Grafakos and T. Tao. Multilinear interpolation between adjoint operators. *J. Funct. Anal.*, 199(2):379–385, 2003.

[14] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.

[15] C. Gu, H. Hu, and Z. Zhou. *Darboux Transformations in Integrable Systems*, volume 26 of *Mathematical Physics Studies*. Springer, Dordrecht, 2005.

[16] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)*, 161(1):397–488, 2005.

[17] M. Lacey. The Bilinear Maximal Functions Map into L^p for $2/3 < p \leq 1$. *Ann. of Math. (2)*, 151(1):35–57, 2000.

[18] M. Lacey and C. Thiele. L^p estimates on the bilinear Hilber transform for $2 < p < \infty$. *Ann. of Math. (2)*, 146(3):693–724, 1997.

[19] M. Lacey and C. Thiele. On Calderón's conjecture. *Ann. of Math. (2)*, 149(2):475–496, 1999.

[20] C. Muscalu. Paraproducts with flag singularities. I. A case study. *Rev. Mat. Iberoam.*, 23(2):705–742, 2007.

[21] C. Muscalu, J. Pipher, T. Tao, and C. Thiele. Bi-parameter paraproducts. *Acta Mathematica*, 193(2):269–296, 2004.

[22] C. Muscalu and W. Schlag. *Classical and Multilinear Harmonic Analysis*, volume 2. Cambridge University Press, to appear.

[23] C. Muscalu and W. Schlag. *Classical and Multilinear Harmonic Analysis*, volume 1. Cambridge University Press, to appear.

[24] C. Muscalu, T. Tao, and C. Thiele. Multi-linear operators given by singular multipliers. *Journal of the American Mathematical Society*, 15(2):469–496, 2002.

[25] C. Muscalu, T. Tao, and C. Thiele. A Counterexample to a Multilinear Endpoint Question of Christ and Kiselev. *Mathematical Research Letters*, 10(2-3):237–246, 2003.

[26] C. Muscalu, T. Tao, and C. Thiele. L^p estimates for the Biest I. The Walsh case. *Math. Ann.*, 329(3):401–426, 2004.

[27] C. Muscalu, T. Tao, and C. Thiele. L^p estimates for the Biest II. The Fourier case. *Math. Ann.*, 329(3):427–461, 2004.

[28] C. Muscalu, T. Tao, and C. Thiele. The Bi-Carleson Operator. *Geom. funct. anal.*, 16(1):230–277, 2006.

[29] R. Oberlin, A. Seeger, T. Tao, C. Thiele, and J. Wright. A variation norm Carleson theorem. *J. Eur. Math. Soc. (JEMS)*, 14(2):421–464, 2012.

[30] E. Palsson. L^p estimates for a singular integral operator motivated by Calderón's second commutator. *J. Funct. Anal.*, 262(4):1645–1678, 2012.

[31] E.M. Stein. On limits of sequences of operators. *Ann. of Math. (2)*, 74:140–170, 1961.

[32] R.S. Strichartz. A multilinear version of the Marcinkiewicz interpolation theorem. *Proc. Amer. Math. Soc.*, 21:441–444, 1969.

[33] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory and Dynamical Systems*, 28(2):657–688, 2008.

[34] C. Thiele. On the Bilinear Hilbert transform. Universität Kiel, Habilitationsschrift, 1998.

[35] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.*, 20(1):53–97, 2007.